Applicants: Graham P. Allaway et al.

Serial No.: 09/888,938 Filed: June 25, 2001

Exhibit 12

### **PCT**

# WORLD INTELLECTUAL PROPERTY ORGANIZATION International Bureau



## INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 5:		(11) International Publication Number: WO 94/22477
A61K 39/12, 39/00, C07K 17/00, 3/00, 13/00, 15/00, C07H 15/12, 13/00	A1	(43) International Publication Date: 13 October 1994 (13.10.94)
(21) International Application Number: PCT/US (22) International Filing Date: 25 March 1994 (20)		European patent (AT, BE, CH, DE, DK, ES, FR, GB, GR
(30) Priority Data: 08/037,816 26 March 1993 (26.03.93)	τ	Published  With international search report.
(60) Parent Application or Grant (63) Related by Continuation US 08/037,8 Filed on 26 March 1993 (2  (71) Applicant (for all designated States except US): PROPHARMACEUTICALS, INC. [US/US]; Old Saw M. Road, Tarrytown, NY 10591 (US).  (72) Inventors; and	26.03.99 GENIC	
<ul> <li>(75) Inventors/Applicants (for US only): HASEL, K [CH/US]; 1713 Baldwin Road, Yorktown Heigi 10598 (US). MADDON, Paul, J. [US/US]; AJ #25C, 60 Haven Avenue, New York, NY 10032 (U</li> <li>(74) Agent: WHITE, John, P.; Cooper &amp; Dunham, 30 Ros Plaza, New York, NY 10112 (US).</li> </ul>	hts, N partmen JS).	

(54) Title: HIV-1 VACCINES, ANTIBODY COMPOSITIONS RELATED THERETO, AND THERAPEUTIC AND PROPHYLACTIC USES THEREOF

#### (57) Abstract

The invention provides a recombinant nucleic acid molecule which encodes a mutant HIV-1 gp120 envelope glycoprotein, vaccines comprising the mutant HIV-1 envelope glycoprotein, antibodies and methods of treating individuals.

Applicants: Graham P. Allaway et al.

Serial No.: 09/888,938 Filed: June 25, 2001

Exhibit 12

## FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	C.D.	** ** * ***		
		GB	United Kingdom	MR	Mau <del>ritania</del>
AU	Australia	GE	Georgia	MW	Malawi
BB	Barbados	GN	Guinea	NE	Niger
BE	Belgium	GR	Greece	NIL	Netherlands
BF	Burkina Feso	EU	Hungary	NO	Norway
BG	Bulgaria	Œ	Ireland	NZ	New Zealand
BJ	Benin	п	Italy	PL.	Poland
BR	Brazil	JP ·	Japan	PT	Portugal
BY	Beiarus	KE	Kenya	RO	Romania
CA	Canada	KG	Kyrgystan	RU	Russian Federation
CF	Central African Republic	KP	Democratic People's Republic	SD	Sudan
CG	Congo		of Kores	SE	Sweden
CH	Switzerland	KR	Republic of Korea	SI	
а	Côte d'Ivoire	KZ.	Kazakhstan	SK	Slovenia
CM	Cameroon	ũ	Liechtenstein		Slovakia
CN	China	LK		SN	Senegal
cs	Czechoslovakia		Sri Lanka	TD	Chad
		LU	Luxenbourg	TG	Togo
CZ	Czech Republic	LV	Latvia	T)	Tajikistan
DΕ	Germany	MC	Monaco	TT	Trinidad and Tobego
DK	Denmark	MID	Republic of Moldova	UA	Ukraine
ES	Spain	MG	Madagascar	US	United States of America
FT	Finland	MIL	Mali	UZ.	Uzbekistan
FR	France	MIN	Mongolia	VN	Viet Nam
GA	Gabon		<u> </u>		

# HIV-1 VACCINES, ANTIBODY COMPOSITIONS RELATED THERETO, AND THERAPEUTIC AND PROPHYLACTIC USES THEREOF

#### Background of the Invention

10

Throughout this application, various publications are referenced by Arabic numerals. Full citations for these references may be found at the end of the specification immediately preceding the claims. The disclosure of these publications is hereby incorporated by reference into this application to describe more fully the art to which this invention pertains.

The life cycle of animal viruses is characterized by a 20 series of events that are required for the productive infection of the host cell. The initial step in the replicative cycle is the attachment of the virus to the cell surface, which attachment is mediated by the specific interaction of the viral attachment protein (VAP) to 25 receptors on the surface of the target cell. differential pattern of expression of these receptors is largely responsible for the host range and tropic properties of viruses. In addition, an effective immune response against many viruses is mediated through neutralizing 30 antibodies directed against the VAP. The interaction of the VAP with cellular receptors and the immune system therefore plays a critical role in infection and pathogenesis of viral disease.

The human immunodeficiency virus type 1 (HIV-1) infects primarily helper T lymphocytes, dendritic cells, and monocytes/macrophages--cells that express surface CD4--leading to a gradual loss of immune function. This loss of function results in the development of the human acquired

2

immunodeficiency syndrome (AIDS) (1). The initial phase of the HIV-1 replicative cycle involves the high-affinity interaction between the HIV-1 exterior envelope glycoprotein gp120 and cell surface CD4 ( $K_a$  approximately 4 x 10.9 M) (2). 5 Several lines of evidence demonstrate the requirement of this interaction for viral infectivity. The introduction into CD4 human cells of cDNA encoding CD4 is sufficient to render otherwise resistant cells susceptible to HIV-1 In vivo, viral infection appears to be infection (3). 10 restricted to cells expressing CD4, indicating that the cellular tropism of HIV-1 is largely determined by the pattern of cellular expression of CD4. Following the binding of HIV-1 gp120 to cell surface CD4, viral and target cell membranes fuse by a mechanism that is poorly 15 understood, resulting in the introduction of the viral capsid into the target cell cytoplasm (4).

Mature CD4 has a relative molecular mass (Mr) of 55 kDa and consists of an N-terminal 372-amino acid extracellular 20 domain containing four tandem immunoglobulin-like regions (V1-V4), followed by a 23-amino acid transmembrane domain and a 38-amino acid cytoplasmic segment (5, 6). experiments using truncated sCD4 proteins, it has been shown that the determinants for high-affinity binding to HIV-1 25 gp120 lie solely within the N-terminal immunoglobulin-like domain (V1) (7-9). Mutational analysis of V1 has defined a discrete binding site (residues 38-52) that comprises a homologous structurally to the second complementarity-determining region (CDR2) of immunoglobulin 30 genes (9).

The production of large quantities of sCD4 has permitted a structural analysis of the two N-terminal immunoglobulin-like domains (V1V2). The structure determined at 2.3

3

angstrom resolution reveals that the molecule has two tightly-associated domains, each of which contains the immunoglobulin-fold connected by a continuous beta strand. The putative binding sites for monoclonal antibodies, class II major histocompatibility complex (MHC) molecules, and HIV-1 gp120, as determined by mutational analyses, map on the molecular surface (10, 11).

The HIV-1 envelope gene env encodes an envelope glycoprotein precursor, gp160, which is cleaved by cellular proteases before transport to the plasma membrane to yield gp120 and gp41. The membrane-spanning glycoprotein, gp41, is non-covalently associated with gp120, a purely extracellular glycoprotein. The mature gp120 molecule is heavily glycosylated (approximately 24 N-linked oligosaccharides), contains approximately 480 amino acid residues with 9 intrachain disulfide bonds (12), and projects from the viral membrane as a dimeric or multimeric molecule (13).

20 Mutational studies of HIV-1 gp120 have delineated important functional regions of the molecule. The regions of gp120 that interact with gp41 map primarily to the N- and Ctermini (14). The predominant strain-specific neutralizing epitope on gp120 is located in the 32-34 amino acid residue 25 third variable loop, herein referred to as the V3 loop, which resides near the center of the gp120 sequence (15). The CD4 binding site maps to discontinuous regions of gp120 that include highly conserved or invariant amino acid residues in the second, third, and fourth conserved domains 30 (the C2, C3, and C4 domains) of gp120 (16). It has been postulated that a small pocket formed by these conserved residues within gp120 could accommodate the CDR2 loop of CD4, a region defined by mutational analyses as important in interacting with gp120 (17).

4

HIV-1 gp120 not only mediates viral attachment to surface CD4 molecules, but also serves as the major target of antibodies which neutralize non-cell-associated virus and inhibit cell to cell viral transmission.

5

There are two major classifications of HIV-1-neutralizing antibodies: type-specific and group-common (15). specific neutralizing antibodies primarily recognize linear determinants in the highly variable V3 loop of gp120. These 10 antibodies act by inhibiting fusion between HIV-1 and the target cell membrane, and generally neutralize only a particular isolate of, or closely related strains of, HIV-1. Sequence variation within the V3 loop, as well as outside of this region, permits viruses to escape neutralization by 15 anti-V3 loop antibodies. In contrast, neutralizing antibodies primarily recognize discontinuous or conformational epitopes in gp120, and possess the ability to neutralize a diverse range of HIV-1 isolates. These broadly neutralizing antibodies often recognize a site on qp120 20 which overlaps the highly conserved CD4-binding site, and thus inhibits gp120-CD4 binding.

A structural relationship has been demonstrated between the V3 loop and the C4 region of gp120 which region constitutes both part of the CD4 binding site and part of the conserved neutralization epitopes. It was observed that deleting the V3 loop resulted in significantly increased binding of a panel of broadly neutralizing hMoAbs (neutralizing human monoclonal antibodies) to the CD4 binding site (18).

30

A major goal in AIDS vaccine development is to develop a vaccine able to protect a subject against the numerous genetic variants of HIV-1 that infect humans. Although cell-mediated immune responses might serve to control infection in HIV-1-infected individuals, several lines of

5

evidence demonstrate that protection against infection is mainly mediated by neutralizing antibodies directed against Early experiments showed that immunization of chimpanzees with recombinant gp120 induced a protective 5 immune response against challenge with the homologous HIV-1 strain (17). This protection correlated with the presence of high-titer neutralizing antibodies against the V3 loop of In addition, passive immunization of chimpanzees with a V3-loop neutralizing monoclonal antibody resulted in 10 protection against challenge with the homologous HIV-1 Although protection against challenge was demonstrated in these two experiments, recent studies have questioned the clinical relevance of these findings. example, these neutralizing antibodies recognize the V3 loop 15 determinants of a single strain, and not conserved or discontinuous epitopes. Thus, these antibodies lack the ability to neutralize the broad spectrum of HIV-1 strains present in an HIV-1 population. Furthermore, the challenge virus was the homologous HIV-1 laboratory adapted LAI (HTLV-20 IIIB) strain and not one of the primary isolates that contain considerable gp120 sequence heterogeneity. these experiments showed that gp120 subunit vaccination induces an immune response effective against only the homogeneous HIV-1 strain used as an antigen, it is unlikely 25 that the vaccination regimens used in these studies would be useful in humans.

Individuals infected by HTV-1 typically develop antibodies that neutralize the virus in vitro, and neutralization titers decrease with disease progression (19). Analysis of sera from HTV-1-infected humans indicates that type-specific neutralizing antibodies appear early in infection. Later in the course of infection, a more broadly neutralizing antibody response develops. However this antibody response is of significantly lower titer and/or affinity.

6

Fractionation studies of HIV-1 antibody-positive human sera reveal that the type-specific neutralizing activity is primarily directed against linear determinants in the V3 loop of gp120 (20). There was no correlation found among antibodies between the ability to neutralize divergent HIV-1 isolates and reactivity to the V3 loop of these isolates. In contrast, the broadly neutralizing antibodies present in HIV-1 antibody-positive human sera primarly recognize discontinuous epitopes in gp120 which overlap the CD4-binding site and block gp120-CD4 binding. In other words, the broadly neutralizing activity of neutralizing antibodies is not merely the result of additive anti-V3 loop reactivities against diverse HIV-1 isolates which appear during infection.

15

Recently, several groups have generated human monoclonal antibodies (hMoAbs) derived from HIV-1 infected individuals which possess type-specific or group-common neutralizing activities (17). The type-specific neutralizing hMoAbs were found to recognize linear determinants in the V3 loop of gp120. In contrast, the group-common neutralizing hMoAbs generally recognize discontinuous epitopes which overlap the CD4-binding site and block gp120-CD4 binding.

The V3 loop is a highly immunodominant region of gp120 which partially interacts with the CD4-binding region. The presence of the V3 loop region on gp120 may skew the humoral immune response away from producing antibodies which specifically bind to the CD4-binding domain of gp120.

Furthermore, the advantages of removing the V3 loop to expose the CD4-binding domain of gp120 to the immune system would be countered by the fact that the exposed CD4-binding site would still have a high affinity for cell surface CD4. In other words, a mutant gp120 protein missing only the V3 loop would quickly bind to CD4+ cells and would thus be

7

hampered in generating an immune response against the exposed CD4-binding site.

The subject invention provides a mutant HIV-1 gp120 envelope

5 glycoprotein which overcomes both the problems of V3 loop immunodominance and of the high affinity to CD4. The subject invention further provides vaccines comprising the mutant HIV-1 gp120 envelope glycoprotein, antibodies which specifically bind to the CD4-binding site of HIV-1 gp120 envelope glycoprotein, pharmaceutical compositions comprising these antibodies, and methods of using these vaccines and compositions to treat or prevent HIV-1 infection.

#### Summary of the Invention

The subject invention provides a recombinant nucleic acid molecule which encodes a mutant HIV-1 gp120 envelope glycoprotein comprising a V3 loop deletion and a C4 domain<sub>(W->X)</sub> point mutation, wherein X is an amino acid residue other than tryptophan. In the preferred embodiment, X is a valine residue.

- 10 In one embodiment, the nucleic acid molecule is a DNA molecule. The DNA molecule may be a plasmid. In one embodiment, the plasmid comprises the sequence of the plasmid designated PPI4-tPA.
- 15 In one embodiment, the C4 domain is an HIV-1<sub>LAI</sub> gp120 envelope glycoprotein C4 domain. The mutant HIV-1 gp120 envelope glycoprotein may be a mutant HIV-1<sub>LAI</sub> gp120 envelope glycoprotein.
- In another embodiment, the C4 domain is an HIV- $1_{\rm R-FL}$  gp120 envelope glycoprotein C4 domain. The mutant HIV-1 gp120 envelope glycoprotein may be a mutant HIV- $1_{\rm R-FL}$  gp120 envelope glycoprotein.
- 25 The subject invention also provides the mutant HIV-1 gp120 envelope glycoprotein encoded by the recombinant nucleic acid molecule of the subject invention.
- The subject invention further provides a vaccine which comprises a therapeutically effective amount of the mutant HIV-1 gp120 envelope glycoprotein of the subject invention, and an adjuvant.

The subject invention further provides a method of treating

9

an HIV-1-infected subject, which comprises immunizing the HIV-1-infected subject with the vaccine of the subject invention, thereby treating the HIV-1-infected subject.

- 5 The subject invention further provides a vaccine which comprises a prophylactically effective amount of the mutant HIV-1 gp120 envelope glycoprotein of the subject invention, and an adjuvant.
- The subject invention further provides a method of reducing the likelihood of an HIV-1-exposed subject's becoming infected with HIV-1, which comprises immunizing the HIV-1-exposed subject with the vaccine of the subject invention, thereby reducing the likelihood of the HIV-1-exposed subject's becoming infected with HIV-1.

The subject invention further provides a method of reducing the likelihood of a non-HIV-1-exposed subject's becoming infected with HIV-1, which comprises immunizing the non-HIV-1-exposed subject with the vaccine of the subject invention, thereby reducing the likelihood of the non-HIV-1-exposed subject's becoming infected with HIV-1.

The subject invention further provides a method of obtaining partially purified antibodies which specifically bind to the CD4-binding domain of HIV-1 gp120 envelope glycoprotein, which method comprises (a) immunizing a non-HIV-1-exposed subject with the vaccine of the subject invention, (b) recovering from the immunized subject serum comprising said antibodies, and (c) partially purifying said antibodies, thereby obtaining partially purified antibodies which specifically bind to the CD4-binding domain of HIV-1 gp120 envelope glycoprotein. In the preferred embodiment, the subject is a human.

10

The subject invention further provides the partially purified antibodies produced by the method of the subject invention.

- 5 The subject invention further provides a pharmaceutical composition, which comprises a therapeutically effective amount of the partially purified antibodies of the subject invention, and a pharmaceutically acceptable carrier.
- The subject invention further provides a method of treating an HIV-1-infected subject, which comprises administering to the subject a dose of the pharmaceutical composition of the subject invention effective to reduce the population of HIV-1-infected cells in the HIV-1-infected subject, thereby treating the HIV-1-infected subject.

The subject invention further provides a method of treating an HIV-1-infected subject, which comprises administering to the subject a dose of the pharmaceutical composition of the subject invention effective to reduce the population of HIV-1 in the HIV-1-infected subject, thereby treating the HIV-1-infected subject.

The subject invention further provides a composition which comprises a prophylactically effective amount of the partially purified antibodies of the subject invention, and a pharmaceutically acceptable carrier.

The subject invention further provides a method of reducing
the likelihood of an HIV-1-exposed subject's becoming
infected with HIV-1, which comprises administering to the
HIV-1-exposed subject a dose of the composition of the
subject invention effective to reduce the population of HIV1 in the HIV-1-exposed subject, thereby reducing the
likelihood of the subject's becoming infected with HIV-1.

11

In one embodiment, the subject is a medical practitioner. In another embodiment, the subject is a newborn infant.

5 Finally, the subject invention provides a method of reducing the likelihood of a non-HIV-1-exposed subject's becoming infected with HIV-1 as a result of exposure thereto during an incident wherein there is an increased risk of exposure to HIV-1, which comprises administering to the subject immediately prior to the incident a dose of the composition of the subject invention effective to reduce the population of HIV-1 to which the subject is exposed during the incident, thereby reducing the likelihood of the subject's becoming infected with HIV-1. In one embodiment, the subject is a medical practitioner.

Brief Description of the Figures

#### Figure 1

gp120 structure. Shown is a box diagram of HIV-1 gp120 depicting the boundaries of the five constant domains (C1-C5) and the five variable domains (V1-V5). The amino acid residue numbering above the box begins at the initiator methionine found at the beginning of the signal sequence (S) and is approximated based on a consensus of all known HIV-1 gp120 amino acid sequences. Also shown are the C4 domain amino acid sequences of HIV-1 strains LAI and JR-FL. Above the C4 domain sequences are indicated two mutations that reduce gp120 binding to cell surface CD4; tryptophan to valine and aspartate to alanine.

12

15

#### Figure 2

PPI4-tPA-gp120LAI. Expression vector with the HIV-1LAI gp120
gene fused to the CMV MIE promoter, and the tPA signal
sequence replacing the HIV-1 gp120 signal sequence.
20 Abbreviations: CMV MIE = cytomegalovirus major immediate
early, E = enhancer, P = promoter, EXA = Exon A, INA =
Intron A, EXB = Exon B, tPA ss = human tissue plasminogen
activator signal sequence, gp120 = glycoprotein 120, BGH =
bovine growth hormone, AMP = ampicillin resistance gene, and
DHFR = dihydrofolate reductase gene.

#### Figure 3

CMV MIE promoter fused to tPA-gp120<sub>LAI</sub>. The nucleotide sequence of the CMV MIE promoter/enhancer region is shown fused to the HIV-1<sub>LAI</sub> gp120 gene that contains the tPA signal sequence. The numbering of nucleotide sequence begins with the HincII site and the numbering of the amino acid sequence begins with the first methionine found in the tPA signal sequence. The tPA signal sequence is fused in-frame to Thr<sub>31</sub>

13

of gp120, the first amino acid found in mature gp120. The signal sequence is shown in bold as are various landmark restriction sites used for cloning as discussed in the text. The locations of Exon A, Intron A, Exon B and the transcription start site and the signal cleavage site are indicated.

#### Figure 4

Transient expression of gp120. Autoradiograph of <sup>35</sup>S-labeled supernatants from COS cell transfectants, immunoprecipitated with a CD4-immunoglobulin-Protein A-Sepharose complex, and run on a reducing 10% SDS-PAGE gel. The plasmids used for transfection were: Lane 1: Mock transfected cells; lane 2: a vector encoding a CD4-immunoglobulin chimera as a positive transfection control; lane 3: PPI4-tPA-gp120<sub>LAI</sub>; and lane 4: PPI4-tPA-gp120<sub>IR-FL</sub>. Positions of molecular weight markers are indicated.

#### Figure 5

Determination of gp120 concentration by ELISA. Panel A: Concentrations of gp120 in media of CHO cell lines, stably transfected with PPI4-tPA-gp120<sub>LAI</sub>, determined by ELISA. Panel B: A standard curve was established using known amounts of gp120.

25

#### Figure 6

Expression of gp120 in stably transfected CHO cells. Autoradiograph of <sup>35</sup>S-labeled supernatants from stable CHO cell lines, immunoprecipitated with a CD4-immunoglobulin30 Protein A-Sepharose complex, and run on a reducing 10% SDS-PAGE gel. Lane 1: clone 9; lane 2: clone 13; lane 3: clone 6; lane 4: Clone 5. Positions of molecular weight markers are indicated.

14

#### Figure 7

tPA-gp120<sub>fR-FL</sub>. The nucleotide and deduced amino acid sequence of the tPA signal sequence fused to HIV-1<sub>fR-FL</sub> gp120 is shown. The NarI and NotI restriction endonuclease sites used for cloning are shown in bold. The predicted site of cleavage by signal peptidase between Arg<sub>35</sub> and Val<sub>36</sub> is indicated.

#### Figure 8

tPA-gp120<sub>LAL</sub>-V3<sup>(c)</sup>. The nucleotide and deduced amino acid sequence of the tPA signal sequence fused to HIV-1<sub>LAI</sub> gp120 with the V3 loop deleted and replaced with the pentapeptide TGAGH is shown. The V3 loop replacement and the NarI and NotI restriction endonuclease sites used for cloning are shown in bold. The predicted site of cleavage by signal peptidase between Arg<sub>35</sub> and Thr<sub>36</sub> is indicated.

#### Figure 9

tPA-gp120<sub>TR-F1</sub>-V3<sup>(\*)</sup>. The nucleotide and deduced amino acid sequence of the tPA signal sequence fused to HIV-1<sub>TR-FL</sub> gp120 with the V3 loop deleted and replaced with the pentapeptide TGAGH is shown. The V3 loop replacement and the NarI and NotI restriction endonuclease sites used for cloning are shown in bold. The predicted site of cleavage by signal peptidase between Arg<sub>35</sub> and Val<sub>36</sub> is indicated.

25

#### Figure 10

tPA-gp120<sub>LAI</sub>-V3<sup>(\*)</sup>-CD4<sup>(\*)</sup>. Shown is the nucleotide and deduced amino acid sequence of the tPA signal sequence fused to HIV-1<sub>LAI</sub> gp120, with the V3 loop deleted and replaced with the pentapeptide TGAGH, and Trp<sub>408</sub> mutated to Val. The mutations and the NarI and NotI restriction endonuclease sites used for cloning are shown in bold. The predicted site of cleavage by signal peptidase between Arg<sub>35</sub> and Thr<sub>36</sub> is indicated.

15

#### Figure 11

tPA-gp120<sub>JR-FL</sub>-V3<sup>(\*)</sup>-CD4<sup>(\*)</sup>. Shown is the nucleotide and deduced amino acid sequence of the tPA signal sequence fused to HIV-l<sub>JR-FL</sub> gp120, with the V3 loop deleted and replaced with the pentapeptide TGAGH, and Trp<sub>3%</sub> mutated to Val. The mutations and the NarI and NotI restriction endonuclease sites used for cloning are shown in bold. The predicted site of cleavage by signal peptidase between Arg<sub>35</sub> and Val<sub>36</sub> is indicated.

10

#### Figure 12

tPA-gp120<sub>LAI</sub>-CD4<sup>(\*)</sup>. Shown is the nucleotide and deduced amino acid sequence of the tPA signal sequence fused to HIV-1<sub>LAI</sub> gp120. The Trp<sub>437</sub> to Val CD4 binding mutation, the NarI and NotI restriction endonuclease sites used for cloning, and the predicted site of cleavage by signal peptidase between Arg<sub>35</sub> and Thr<sub>36</sub> are shown in bold.

#### Figure 13

tPA-gpl20<sub>IR.FL</sub>-CD4<sup>(\*)</sup>. Shown is the nucleotide and deduced amino acid sequence of the tPA signal sequence fused to HIV-1<sub>IR.FL</sub> gpl20. The Trp<sub>424</sub> to Val CD4 binding mutation, the NarI and NotI restriction endonuclease sites used for cloning and the predicted cleavage by signal peptidase between Arg<sub>35</sub> and Val<sub>36</sub> are shown in bold.

#### Figure 14

Expression of gp120 in stably transfected CHO cells.

Autoradiograph of super <sup>35</sup>S-labeled supernatants from stable CHO cell lines, immunoprecipitated with MoAb F105-Protein A-Sepharose complex, and run on a reducing 10% SDS-PAGE gel. Panel A: Lane 1: tPA-gp120<sub>LAI</sub> CHO cells; lane 2: tPA-gp120<sub>LAI</sub>-V3<sup>(·)</sup> CHO cells; lane 3: tPA-gp120<sub>LAI</sub>-V3<sup>(·)</sup>-CD4<sup>(·)</sup> CHO cells. Panel B: Lane 1: tPA-gp120<sub>IR-FL</sub> CHO cells; lane 2: tPA-gp120<sub>IR-FL</sub>-V3<sup>(·)</sup>

16

CHO cells; lane 3: tPA-gp120<sub>JR-FL</sub>-V3<sup>(-)</sup>-CD4<sup>(-)</sup> CHO cells. Positions of molecular weight markers are indicated.

#### Figure 15

#### 5 Purified ap120 proteins.

Silver stained 10% SDS-PAGE gel with a sample of purified gp120 proteins. Panel A: Lane 1: tPA-gp120<sub>LAI</sub> CHO cells; lane 2: tPA-gp120<sub>LAI</sub>-V3<sup>(·)</sup> CHO cells; lane 3: tPA-gp120<sub>LAI</sub>-V3<sup>(·)</sup>-CD4<sup>(·)</sup> CHO cells. Panel B: Lane 1: tPA-gp120<sub>IR-FL</sub> CHO cells; lane 2: tPA-gp120<sub>IR-FL</sub>-V3<sup>(·)</sup>-CD4<sup>(·)</sup> CHO cells. Positions of molecular weight markers are indicated.

#### Figure 16

Analysis of binding of recombinant mutant gp120 to cell
15 surface human CD4 by FACS.

Plate 1. DG44 cells, a subclone of CHO cells which lack expression of the human CD4 protein, were used as control. Increasing concentrations of HIV-1  $gp120_{LAI}$  did not show an specific fluoresence when compared to increase in Plate 2. DG44 #3 cells are a CHO cell line 20 background. transfected with the cDNA clone encoding the human CD4 protein. Increasing concentrations of HIV-1 gp120[A] show a dramatic increase (or shift) in fluoresence. Similar to Plate 2 but the HIV-1 gp120<sub>LAI</sub>-V3<sup>(-)</sup> protein was 25 added. Again a large shift indicating binding to the DG44 #3 cells was seen. Plate 4. DG44 #3 cells were incubated with either HIV-1 gpl20<sub>LAI</sub>-V3<sup>(-)</sup>-CD4<sup>(-)</sup> protein or MoAb OKT4A an antibody with high affinity for human CD4. Only OKT4A bound to the cells.

### Detailed Description of the Invention

The plasmids designated PPI4-tPA-gp120<sub>LAI</sub> and PPI4-tPA-gp120<sub>JR</sub>

fl were deposited pursuant to, and in satisfaction of, the

requirements of the Budapest Treaty on the International Recognition of the Deposit of Microorganisms for the Purposes of Patent Procedure with the American Type Culture Collection (ATCC), 12301 Parklawn Drive, Rockville, Maryland 20852 under ATCC Accession Nos. 75431 and 75432, respectively. The plasmids PPI4-tPA-gp120<sub>LAI</sub> and PPI4-tPA-gp120<sub>IR-FL</sub> were deposited with the ATCC on March 12, 1993.

The subject invention provides a recombinant nucleic acid molecule which encodes a mutant HIV-1 gp120 envelope glycoprotein comprising a V3 loop deletion and a C4 domain<sub>(W->X)</sub> point mutation, wherein X is an amino acid residue other than tryptophan. In the preferred embodiment, X is a valine residue.

- In one embodiment, the nucleic acid molecule is a DNA molecule. The DNA molecule may be a plasmid. In one embodiment, the plasmid comprises the sequence of the plasmid designated PPI4-tPA.
- The V3 loop of HIV-1 gp120 envelope glycoprotein is shown in Figure 1. The V3 loop is demarcated by cysteine residues at both its N- and C-termini. As used herein, a V3 loop deletion means a deletion of one or more amino acid residues between the terminal cysteine residues, with the proviso that there must be three or more amino acid residues situated between the two terminal cysteine residues in a V3 loop deletion. These three or more amino acid residues may either be residues originally present in the V3 loop, or exogenous residues. For example, as shown in the

18

Experimental Details section <u>infra</u>, the pentapeptide TGAGH is situated between the two terminal cysteine residues. Variations in the size of the V3 loop deletion illustrated herein are tolerable without affecting the overall structure of the mutant HIV-1 gp120 envelope glycoprotein, as is well known to those skilled in the art.

As used herein, "C4 domain" means the HIV-1 gp120 envelope glycoprotein C4 domain having the following consensus 10 sequence:

 $X_{1}X_{2}X_{3}CX_{4}IX_{5}X_{6}X_{7}X_{8}X_{9}X_{10}WX_{11}X_{12}X_{13}X_{14}X_{15}AX_{16}YX_{17}X_{18}-PX_{19}X_{20}X_{21}X_{22}X_{23}X_{24}X_{25}X_{26}SX_{77}X_{28}TGX_{29}X_{30}X_{31}X_{32}RX_{33}GX_{34}$ 

15 wherein X<sub>1</sub> = T, I, V, K or R; X<sub>2</sub> = L, I or H; X<sub>3</sub> = P, Q, L or
T; X<sub>4</sub> = R, K or G; X<sub>5</sub> = K or E; X<sub>6</sub> = Q or E; X<sub>7</sub> = F, I or V;
X<sub>8</sub> = I, V or M; X<sub>9</sub> = N, R or K; X<sub>10</sub> = M, R, L or T; X<sub>11</sub> = Q, R
or V; X<sub>12</sub> = E, K, G, R, V or A; X<sub>13</sub> = V, T, A or G; X<sub>14</sub> = G or
E; X<sub>15</sub> = K, R, E, or Q; X<sub>16</sub> = M, V, I or L; X<sub>17</sub> = A, T or D; X<sub>18</sub>
20 = P or L; X<sub>19</sub> = I or F; X<sub>20</sub> = S, R, G, K, N, A, E or Q; X<sub>21</sub> =
G or R; X<sub>22</sub> = Q, L, P, N, K, V, T, E or I; X<sub>23</sub> = I, V or L; X<sub>24</sub>
= R, K, S, N, G, I, T, E or I; X<sub>25</sub> = C or R; X<sub>26</sub> = S, L, I, T,
P, E, V, K, D or N; X<sub>27</sub> = N, K or L; X<sub>28</sub> = I or V; X<sub>29</sub> = L, P
or I; X<sub>30</sub> = L or I; X<sub>31</sub> = L or I; X<sub>32</sub> = T, A, I, V or E; X<sub>33</sub> =
25 D or E; X<sub>34</sub> = G or V.

The C4 domain consensus sequence is based on existing C4 domain sequence information from various HIV-1 strains, and thus is not necessarily an exhaustive consensus sequence. The conserved tryptophan residue shown in bold after residue  $X_{10}$  is the only conserved tryptophan residue in the C4 domain. As used herein, a C4 domain $_{(W->X)}$  point mutation is a mutation of the above-identified conserved C4 domain tryptophan residue to an amino acid residue other than

19

tryptophan. For example, a C4 domain  $(w_->v)$  point mutation is a mutation of the conserved C4 domain tryptophan residue to a valine residue.

In one embodiment, the C4 domain is an HIV-1<sub>LAI</sub> gp120 envelope glycoprotein C4 domain. The sequence of the HIV-1<sub>LAI</sub> gp120 C4 domain is: TLPCRIKQFINMWQEVGKAMYAPPISGQIRCS-SNITGLLLTRDGG. The mutant HIV-1 gp120 envelope glycoprotein may be a mutant HIV-1<sub>LAI</sub> gp120 envelope glycoprotein.

10

In another embodiment, the C4 domain is an HIV-1<sub>IR-FL</sub> gp120 envelope glycoprotein C4 domain. The sequence of the HIV-1<sub>IR-FL</sub> gp120 C4 domain is: TLPCRIKQIINMWQEVGKAMYAPPIRGQIRCS-SNITGLLLTRDGG. The mutant HTV-1 gp120 envelope glycoprotein may be a mutant HIV-1<sub>IR-FL</sub> gp120 envelope glycoprotein.

HIV-1<sub>LAI</sub> is a laboratory-adapted strain that is tropic for phytohemagglutinin (PHA)-stimulated peripheral lymphocytes (PBLs) and immortalized human T-cell lines. In 20 contrast, HIV-1<sub>JR-FL</sub> was isolated from brain tissue taken at autopsy that was co-cultured with lectin-activated normal human PBLs.  $HIV-1_{JR-FL}$  is tropic for PHA-stimulated PBLs and blood-derived macrophages but will not replicate in transformed T-cell lines. Mutant HIV-1 gp120 envelope 25 glycoproteins derived from a clinical isolate of HIV-1 such as JR-FL may possess new or different epitopes compared to the laboratory-adapted HIV-1 strains that are beneficial for successful vaccination. Although only the HIV-1LM and HIV- $1_{\text{R-PL}}$  strains are used herein to generate the mutant HIV-1 30 gp120 envelope glycoproteins of the subject invention, other HIV-1 strain could be substituted in their place as is well.

The V1 and V2 variable regions of gp120 are unnecessary for

known to those skilled in the art.

WO 94/22477

20

PCT/US94/03282

CD4 binding (21). Therefore the mutant HIV-1 gp120 envelope glycoprotein of this invention can either include or exclude the V1 and V2 variable regions.

The subject invention additionally provides a recombinant nucleic acid molecule which encodes a mutant HIV-1 gp120 envelope glycoprotein comprising a V3 loop deletion and a C4 domain<sub>(Asp->X)</sub> point mutation, wherein the aspartate residue is between amino acid residues X<sub>15</sub> and X<sub>16</sub> in the C4 consensus sequence, and X is an amino acid residue other than aspartate or glutamate. In the preferred embodiment, X is an alanine residue.

The subject invention additionally provides a recombinant nucleic acid molecule which encodes a mutant HIV-1 gp120 envelope glycoprotein comprising a V3 loop deletion and a C4 domain<sub>(Gh->X)</sub> point mutation, wherein the glutamate residue is between amino acid residues X<sub>15</sub> and X<sub>16</sub> in the C4 consensus sequence, and X is an amino acid residue other than aspartate or glutamate. In the preferred embodiment, X is an alanine residue.

The subject invention additionally provides a recombinant nucleic acid molecule which encodes a mutant HIV-1<sub>LAI</sub> gp120 envelope glycoprotein comprising a V3 loop deletion and a C3 domain<sub>(exp378->X)</sub> point mutation, wherein X is an amino acid residue other than aspartate or glutamate. In the preferred embodiment, X is a lysine residue.

The subject invention additionally provides a recombinant nucleic acid molecule which encodes a mutant HIV-1<sub>R-FL</sub> gp120 envelope glycoprotein comprising a V3 loop deletion and a C3 domain<sub>(ap369->X)</sub> point mutation, wherein X is an amino acid residue other than aspartate or glutamate. In the preferred

21

embodiment, X is a lysine residue.

The subject invention additionally provides a recombinant nucleic acid molecule which encodes a mutant HIV-1<sub>LAI</sub> gp120 envelope glycoprotein comprising a V3 loop deletion and: a C3 domain<sub>(gh380->X)</sub> point mutation, wherein X is an amino acid residue other than glutamate. In the preferred embodiment, X is a glutamine residue.

The subject invention additionally provides a recombinant nucleic acid molecule which encodes a mutant HIV-1<sub>JR-FL</sub> gp120 envelope glycoprotein comprising a V3 loop deletion and a C3 domain<sub>(ghi71->X)</sub> point mutation, wherein X is an amino acid residue other than glutamate. In the preferred embodiment,

15 X is a glutamine residue.

The subject invention additionally provides a recombinant nucleic acid molecule which encodes a mutant HIV-1<sub>LAI</sub> gp120 envelope glycoprotein comprising a V3 loop deletion and a C2 domain<sub>(thr267->X)</sub> point mutation, wherein X is an amino acid residue other than threonine. In the preferred embodiment, X is an arginine residue.

The subject invention additionally provides a recombinant nucleic acid molecule which encodes a mutant HTV-1<sub>JR-FL</sub> gp120 envelope glycoprotein comprising a V3 loop deletion and a C2 domain<sub>(th/260->X)</sub> point mutation, wherein X is an amino acid residue other than threonine. In the preferred embodiment, X is an arginine residue.

30

The subject invention additionally provides a recombinant nucleic acid molecule which encodes a mutant HIV-1 gp120 envelope glycoprotein comprising (a) a V3 loop deletion, or (b) a one of the C2, C3 or C4 domain point mutations

WO 94/22477

discussed supra.

The point mutations in the recombinant nucleic acid molecules described <u>supra</u> are selected based on their ability to reduce the affinity of the mutant gp120 glycoprotein encoded thereby for CD4. As used herein, the term "reduce the affinity" means to reduce the affinity by at least two-fold.

PCT/US94/03282

One skilled in the art would know how to make recombinant nucleic acid molecules which encode mutant HIV-1 gp120 envelope glycoproteins comprising a V3 loop deletion and the specific C2, C3 or C4 domain point mutations corresponding to those mutations exemplified in the HIV-1<sub>IR-FL</sub> and HIV-1<sub>IA</sub> strains, supra. Furthermore, one skilled in the art would know how to use these recombinant nucleic acid molecules to obtain the proteins encoded thereby, and practice the therapeutic and prophylactic methods of using same, as described herein for the recombinant nucleic acid molecule which encodes a mutant HIV-1 gp120 envelope glycoprotein comprising a V3 loop deletion and a C4 domain<sub>(W->X)</sub> point mutation.

The subject invention also provides the mutant HIV-1 gp120 envelope glycoprotein encoded by the recombinant nucleic acid molecule of the subject invention.

In accordance with the invention, numerous vector systems for expression of the mutant HIV-1 gp120 envelope glycoprotein may be employed. For example, one class of vectors utilizes DNA elements which are derived from animal viruses such as bovine papilloma virus, polyoma virus, adenovirus, vaccinia virus, baculovirus, retroviruses (RSV, MMTV or MoMLV), Semliki Forest virus or SV40 virus.

23

Additionally, cells which have stably integrated the DNA into their chromosomes may be selected by introducing one or more markers which allow for the selection of transfected host cells. The marker may provide, for example, prototropy auxotrophic host, biocide resistance, antibiotics) or resistance to heavy metals such as copper or the like. The selectable marker gene can be either directly linked to the DNA sequences to be expressed, or introduced into the same cell by cotransformation. Additional elements 10 may also be needed for optimal synthesis of mRNA. These include splice elements may signals, transcriptional promoters, enhancers, and termination The cDNA expression vectors incorporating such elements include those described by Okayama (22).

15

The vectors used in the subject invention are designed to express high levels of mutant HIV-1 gp120 envelope glycoproteins in cultured eukaryotic cells as well efficiently secrete these proteins into the culture medium. 20 The targeting of the mutant HIV-1 gp120 glycoproteins into the culture medium is accomplished by fusing in-frame to the mature N-terminus of the mutant HIV-1 gp120 envelope glycoprotein the tissue plasminogen activator (tPA) prepro-signal sequence.

25

The mutant HIV-1 gp120 envelope glycoprotein may be produced by a) transfecting a mammalian cell with an expression vector producing HIV-1 for mutant gp120 envelope glycoprotein; b) culturing the resulting transfected mammalian cell under conditions such that mutant HIV-1 gp120 envelope glycoprotein is produced; and c) recovering the mutant HIV-1 gp120 envelope glycoprotein so produced.

Once the expression vector or DNA sequence containing the constructs has been prepared for expression, the expression

production of the mutant glycoprotein.

vectors may be transfected or introduced into an appropriate mammalian cell host. Various techniques may be employed to achieve this, such as, for example, protoplast fusion, calcium phosphate precipitation, electroporation or other conventional techniques. In the case of protoplast fusion, the cells are grown in media and screened for the appropriate activity. Expression of the gene encoding a mutant HIV-1 gp120 envelope glycoprotein results in

24

10

Methods and conditions for culturing the resulting transfected cells and for recovering the mutant HIV-1 gp120 envelope glycoprotein so produced are well known to those skilled in the art, and may be varied or optimized depending upon the specific expression vector and mammalian host cell employed.

In accordance with the claimed invention, the preferred host cells for expressing the mutant HTV-1 gp120 envelope glycoprotein of this invention are mammalian cell lines. Mammalian cell lines include, for example, monkey kidney CV1 line transformed by SV40 (COS-7); human embryonic kidney line 293; baby hamster kidney cells (BHK); Chinese hamster ovary-cells-DHFR (CHO); Chinese hamster ovary-cells DHFR (DXB11); monkey kidney cells (CV1); African green monkey kidney cells (VERO-76); human cervical carcinoma cells (HELA); canine kidney cells (MDCK); human lung cells (W138); human liver cells (Hep G2); mouse mammary tumor (MMT 060562); mouse cell line (C127); and myeloma cell lines.

30

Other eukaryotic expression systems utilizing non-mammalian vector/cell line combinations can be used to produce the mutant HIV-1 gp120 envelope glycoproteins. These include, but are not limited to, baculovirus vector/insect cell

expression systems and yeast shuttle vector/yeast cell expression systems.

25

- Methods and conditions for purifying mutant HIV-1 gp120 envelope glycoproteins from the culture media are provided in the invention, but it should be recognized that these procedures can be varied or optimized as is well known to those skilled in the art.
- The subject invention further provides a vaccine which comprises a therapeutically effective amount of the mutant HIV-1 gp120 envelope glycoprotein of the subject invention, and an adjuvant.
- 15 A therapeutically effective amount of the mutant HIV-1 gp120 envelope glycoprotein may be determined according to methods well known to those skilled in the art.
- As used herein, adjuvants include, but are not limited to, alum, Freund's incomplete adjuvant (FIA), Saponin, Quil A, Monophosphoryl lipid A (MPL), and nonionic block copolymers (SAF) such as L-121 (Pluronic; Syntex SAF). In the preferred embodiment, the adjuvant is alum, especially in the form of a thixotropic, viscous, and homogeneous aluminum hydroxide gel. The vaccine of the subject invention may be administered as an oil in water emulsion. Methods of combining adjuvants with antigens are well known to those skilled in the art.
- 30 The subject invention further provides a method of treating an HIV-1-infected subject, which comprises immunizing the HIV-1-infected subject with the vaccine of the subject invention, thereby treating the HIV-1-infected subject.
- 35 As used herein, treating an HIV-1-infected subject with the

PCT/US94/03282

vaccine of the subject invention means reducing in the subject either the population of HIV-1 or HIV-1-infected cells, or ameliorating the progression of an HIV-1-related disorder in the subject.

5

As used herein, an "HIV-infected subject" means an individual having at least one of his own cells invaded by HIV-1.

- 10 As used herein, "immunizing" means administering a primary dose of the vaccine to a subject, followed after a suitable period of time by one or more subsequent administrations of the vaccine, so as to generate in the subject an immune response against the CD4-binding region of the mutant HIV-1 gp120 envelope glycoprotein in the vaccine. A suitable period of time between administrations of the vaccine may readily be determined by one skilled in the art, and is usually in the order of several weeks to months.
- 20 In the preferred embodiment, the dose of vaccine administered is an amount sufficient to deliver to the subject between 10ug and 1mg of the mutant HIV-1 gp120 envelope glycoprotein.
- The subject invention further provides a vaccine which comprises a prophylactically effective amount of the mutant HIV-1 gp120 envelope glycoprotein of the subject invention, and an adjuvant.
- 30 A prophylactically effective amount of the mutant HIV-1 gp120 envelope glycoprotein may be determined according to methods well known to those skilled in the art.
- The subject invention further provides a method of reducing the likelihood of an HIV-1-exposed subject's becoming

27

infected with HIV-1, which comprises immunizing the HIV-1-exposed subject with the vaccine of the subject invention, thereby reducing the likelihood of the HIV-1-exposed subject's becoming infected with HIV-1.

5

As used herein, the subject's becoming infected with HIV-1 means the invasion of the subject's own cells by HIV-1.

As used herein, reducing the likelihood of a subject's becoming infected with HIV-1 means reducing the likelihood of the subject's becoming infected with HIV-1 by at least two-fold. For example, if a subject has a 1% chance of becoming infected with HIV-1, a two-fold reduction in the likelihood of the subject's becoming infected with HIV-1 would result in the subject's having a 0.5% chance of becoming infected with HIV-1. In the preferred embodiment of this invention, reducing the likelihood of the subject's becoming infected with HIV-1 means reducing the likelihood of the subject's becoming infected with HIV-1 by at least ten-fold.

As used herein, an HIV-1-exposed subject is a subject who has HIV-1 present in his body, but has not yet become HIV-1-infected.

25

The subject invention further provides a method of reducing the likelihood of a non-HIV-1-exposed subject's becoming infected with HIV-1, which comprises immunizing the non-HIV-1-exposed subject with the vaccine of the subject invention, thereby reducing the likelihood of the non-HIV-1-exposed subject's becoming infected with HIV-1.

As used herein, a non-HIV-1-exposed subject is a subject who does not have HIV-1 present in his body.

The subject invention further provides a method of obtaining partially purified antibodies which specifically bind to the CD4-binding domain of HIV-1 gp120 envelope glycoprotein, which method comprises (a) immunizing a non-HIV-1-exposed 5 subject with the vaccine of the subject invention, recovering from the immunized subject serum comprising said antibodies, and (c) partially purifying said antibodies, thereby obtaining partially purified antibodies which specifically bind to the CD4-binding domain of HIV-1 qp120 In the preferred embodiment, the 10 envelope glycoprotein. subject is a human.

PCT/US94/03282

As used herein, partially purified antibodies means a composition which comprises antibodies which specifically 15 bind to the CD4-binding domain of HIV-1 gp120 envelope glycoprotein, and consists of fewer protein impurities than does the serum from which the anti-CD4-binding domain antibodies are derived. A protein impurity means a protein other than the anti-CD4-binding domain antibodies. example, the partially purified antibodies might be an IqG preparation.

Methods of recovering serum from a subject are well known to those skilled in the art. Methods of partially purifying 25 antibodies are also well known to those skilled in the art, and include, by way of example, filtration, ion exchange chromatography, and precipitation.

In one embodiment, the partially purified antibodies 30 comprise an immune globulin (IG) preparation. IG can be purified from serum by a two-step process. Initially, serum is fractionated by the cold ethanol method of Cohn, et al. Cohn Fraction II has as its main protein component immunoglobulin present as monomers, dimers 35 aggregates. Fraction II is then purified to produce IVIG

29

(immune globulin intravenous) using a variety of purification methods which include, for example, ion exchange, DEAE chromatography, acid pH 4.25 diafiltration, PEG precipitation or Pepsin treatment. The final product is stabilized (e.g., glucose + NaCl) and the final IgG concentration is fixed at between about 3% and about 6%.

The subject invention further provides the partially purified antibodies produced by the method of the subject invention.

The subject invention further provides a pharmaceutical composition, which comprises a therapeutically effective amount of the partially purified antibodies of the subject invention, and a pharmaceutically acceptable carrier.

A therapeutically effective amount of the partially purified antibodies of the subject invention may be determined according to methods well known to those skilled in the art.

20

10

Pharmaceutically acceptable carriers are well known to those skilled in the art and include, but are not limited to, 0.01-0.1M and preferably 0.05M phosphate buffer or 0.8% Additionally, such pharmaceutically acceptable saline. 25 carriers may be aqueous or non-aqueous solutions, suspensions, and emulsions. Examples of non-aqueous solvents are propylene glycol, polyethylene vegetable oils such as olive oil, and injectable organic esters such as ethyl oleate. Aqueous carriers include alcoholic/aqueous 30 water, solutions, emulsions suspensions, including saline and buffered media. Parenteral vehicles include sodium chloride solution, Ringer's dextrose, dextrose and sodium chloride, lactated Ringer's or fixed oils. Intravenous vehicles include fluid 35 and nutrient replenishers, electrolyte replenishers such as

those based on Ringer's dextrose, and the like. Preservatives and other additives may also be present, such as, for example, antimicrobials, antioxidants, chelating agents, inert gases and the like.

30

5

The subject invention further provides a method of treating an HIV-1-infected subject, which comprises administering to the subject a dose of the pharmaceutical composition of the subject invention effective to reduce the population of HIV-1-infected cells in the HIV-1-infected subject, thereby treating the HIV-1-infected subject.

As used herein, administering may be effected or performed using any of the various methods known to those skilled in the art. The administering may comprise administering intravenously. The administering may also comprise administering intramuscularly. The administering may further comprise administering subcutaneously.

The dose of the pharmaceutical composition of the subject invention effective to reduce the population of HIV-1-infected cells in the HIV-1-infected subject may be readily determined using methods well known to those skilled in the art. In the preferred embodiment, the dose is sufficient to deliver to the subject between about 10 mg/kg and 150mg/kg of protein if administered intramuscularly. In the preferred embodiment, the dose is sufficient to deliver to the subject between about 100 mg/kg and 2g/kg of protein if administered intravenously.

30

The subject invention further provides a method of treating an HIV-1-infected subject, which comprises administering to the subject a dose of the pharmaceutical composition of the subject invention effective to reduce the population of HIV-1 in the HIV-1-infected subject, thereby treating the HIV-1-

31

infected subject.

The dose of the pharmaceutical composition of the subject invention effective to reduce the population of HIV-1 in the HIV-1-infected subject may be readily determined using methods well known to those skilled in the art. In the preferred embodiment, the dose is sufficient to deliver to the subject between about 10 mg/kg and 150mg/kg of protein if administered intramuscularly. In the preferred embodiment, the dose is sufficient to deliver to the subject between about 100 mg/kg and 2g/kg of protein if administered intravenously.

The subject invention further provides a composition which
comprises a prophylactically effective amount of the
partially purified antibodies of the subject invention, and
a pharmaceutically acceptable carrier.

A prophylactically effective amount of the partially purified antibodies of the subject invention may be determined according to methods well known to those skilled in the art.

The subject invention further provides a method of reducing
the likelihood of an HIV-1-exposed subject's becoming
infected with HIV-1, which comprises administering to the
HIV-1-exposed subject a dose of the composition of the
subject invention effective to reduce the population of HIV1 in the HIV-1-exposed subject, thereby reducing the
likelihood of the subject's becoming infected with HIV-1.

In one embodiment, the subject is a medical practitioner. The medical practitioner may be a medical practitioner exposed to an HTV-1-containing bodily fluid. As used herein, the term "medical practitioner" includes, but is in no way

32

limited to, doctors, dentists, surgeons, nurses, medical laboratory assistants, and students in health care programs.

In another embodiment, the subject is a newborn infant. The newborn infant may be a newborn infant born to an HIV-1-infected mother.

The dose of the composition of the subject invention effective to reduce the population of HIV-1 in the HIV-110 exposed subject may be readily determined using methods well known to those skilled in the art. In the preferred embodiment, the dose is sufficient to deliver to the subject between about 10mg/kg and 150mg/kg of protein if administered intramuscularly. In the preferred embodiment, the dose is sufficient to deliver to the subject between about 100 mg/kg and 2g/kg of protein if administered intravenously.

The vaccines and pharmaceutical compositions of the subject invention may also ameliorate the progression of an HIV-1-related disorder in a subject to whom the vaccines or pharmaceutical compositions were administered while the subject was either non-HIV-1-exposed or HIV-1-exposed, but not yet HIV-1-infected.

25

Finally, the subject invention provides a method of reducing the likelihood of a non-HIV-1-exposed subject's becoming infected with HIV-1 as a result of exposure thereto during an incident wherein there is an increased risk of exposure to HIV-1, which comprises administering to the subject immediately prior to the incident a dose of the composition of the subject invention effective to reduce the population of HIV-1 to which the subject is exposed during the incident, thereby reducing the likelihood of the subject's becoming infected with HIV-1. In one embodiment, the

33

subject is a medical practitioner.

An incident wherein there is an increased risk of exposure to HIV-1 includes, for example, receiving a blood transfusion, sexual contact with an HIV-1-infected individual, and performing a HIV-1-containing bodily fluid-exposing medical procedure.

As used herein, "immediately prior to the incident" means 10 within one month of the incident. In the preferred embodiment, "immediately prior to the incident" means within one day of the incident.

The dose of the composition of the subject invention effective to reduce the population of HIV-1 to which the subject is exposed during the incident may be readily determined using methods well known to those skilled in the art. In the preferred embodiment, the dose is sufficient to deliver to the subject between about 10mg/kg and 150mg/kg of protein if administered intramuscularly. In the preferred embodiment, the dose is sufficient to deliver to the subject between about 100mg/kg and 2g/kg of protein if administered intravenously.

of this invention is a method embodiment 25 substantially reducing the likelihood of a non-infected medical practitioner's becoming infected with HIV-1 during a bodily fluid-exposing medical procedure involving a patient, which comprises administering to the patient during a suitable time period an amount of the composition of the 30 subject invention effective to substantially reduce the likelihood of the non-infected medical practitioner's becoming infected with HIV-1 by virtue of contact with the patient's bodily fluid during the medical procedure.

As used herein, a bodily fluid is any fluid which is present in the human body and is capable of containing infectious HIV-1 in an HIV-1-infected patient. Bodily fluids include, but are not limited to, saliva, cerebrospinal fluid, tears, vaginal secretions, urine, alveolar fluid, synovial fluid and pleural fluid.

Another embodiment of this invention is a method of substantially reducing the likelihood of a non-HIV-110 infected newborn infant's becoming infected with HIV-1 prior to or during birth from an HIV-1-infected mother, which comprises administering to the mother prior to birth an amount of the composition of the subject invention effective to substantially reduce the likelihood of the non-HIV-115 infected newborn infant's becoming infected with HIV-1 by virtue of contact with the patient's bodily fluid.

In order to facilitate an understanding of the Experimental Details section which follows, certain frequently occurring 20 methods and/or terms are best described in Maniatis et al. (23).

This invention will be better understood by reference to the Experimental Details which follow, but those skilled in the art will readily appreciate that the specific experiments detailed are only illustrative of the invention as described more fully in the claims which follow thereafter.

# Experimental Details

### Nomenclature

As used herein, V3<sup>(·)</sup> indicates a V3 loop deletion from HIV-1 gp120 envelope glycoprotein. As used herein, CD4<sup>(·)</sup> indicates a point mutation in the C4 domain of HIV-1 gp120 envelope glycoprotein which mutation inhibits CD4 binding to the mutant HIV-1 gp120 envelope glycoprotein. The structure of HIV-1 gp120 envelope glycoprotein is shown in Figure 1.

### 10 Materials and Methods

Construction of PPI4-tPA-gp120, expression vector. 1. An expression vector was constructed that consisted of the cytomegalovirus major immediate early (CMV promoter/enhancer linked to the HIV-1 env gene, which gene 15 had its signal sequence replaced by the tPA signal sequence. The CMV MIE promoter/enhancer sequences were derived from pSVCC1 (24) consisting of 1580 base pairs of contiguous DNA that is immediately 5' to the initiator ATG. In sequential 20 order, the functional domains of the CMV promoter are: the promoter/enhancer region; a transcriptional initiator site; exon A (a non-coding exon); intron A; and 17 nucleotides of exon B (non-coding sequences). The viral promoter sequences were ligated to a gene construct consisting of 25 nucleotide sequences encoding amino acids -35 to -1 of human tPA (25) fused in-frame to HIV-1 amino acids 31 through 515, ending with a TGA stop codon. The construction was performed in two parts. The majority of the CMV promoter could be isolated as a 1560 bp Hinc II/Pst I fragment which 30 was ligated to a Pst I/Not I 1590 bp DNA fragment that contained the remainder of the CMV promoter, the initiator ATG, the tPA signal sequence and the mature HIV-11AI env protein coding sequence.

30

The latter fragment was assembled using the polymerase chain reaction as follows. Primer 1 (GATCCTGCAGTCACCGTCCTTGACA-CGATGGATGCAATGAAGAGA) and primer 2 (AAGTCTTCTCCTCGGTCTTGT-CTTTTTAACACCCAG) were used to amplify the nucleic acid 5 sequences encoding the tPA signal sequence amino acids -35 to -1 from plasmid pMAM neo-s (Clonetech), thus producing a 150 bp fragment. A second 1440 bp DNA fragment was amplified (TTCAGAAGAGGAGCCAGAACAGAAAATTGTGGGTC), primer 3 primer 4 (GGAAAAAGCGGCCGCTCATTTTTCTCTCTGCACCACTC), and DENV The PCR fragments were pooled, (26) as a template. 10 desalted, and excess primer removed by ultrafiltration through a centricon-100 unit (Amicon). An aliquot of the pooled material was then subjected to a second round of amplification in the presence of primers 1 and 4 to produce a 1590 bp fragment, which was then digested with Pst I and The CMV promoter fragment and the HIV-1, an env ligated together, and the entire fragment were then transcription unit subcloned into PPI4, which eukaryotic shuttle vector that contains an ampicillin 20 resistance gene, an SV40 origin of replication and a DHFR gene whose transcription is driven by the ß-globin promoter. The final construct, PPI4-tPA-gp120LAI, is shown in Figure 2.

The expression vector is then used as the prototype vector 25 for the expression of gp120 proteins that are derived from other HIV-1 strains or mutated as described in the methods The vector was constructed so that unique Nar I and Not I sites flank the gp120 sequence, thus facilitating the removal of the gp120 gene cassette and the subsequent insertion of other gene cassettes (Figure 2).

- Expression of HIV-1, a gp120 in mammalian cells. 2.
- Transient expression.

CosM5 cells grown in DMEM containing 10% fetal calf serum

25

were split to 75% confluence. On the following day, the cells were transfected for 16-20 hours with 10 micrograms of CsCl-purified PPI4-tPA-gp120<sub>LAI</sub> DNA by the standard CaPO<sub>4</sub> (5) precipitation technique. After transfection, fresh medium added to the cells. Analysis of the products synthesized 96-120 hours post-transfection was performed by radiolabelling the transfectants with 35S-cysteine for 12-18 hours, followed by precipitation of media using a CD4immunoglobulin-Protein A-Sepharose complex, followed by SDS-10 PAGE under reducing conditions (Figure 4).

## Stable expression.

Dhfr Chinese hamster ovary cells (CHO) were transfected with 20 micrograms of CsCl-purified DNA. Approximately 3-5 15 days post-transfection, cells were placed in selective medium (nucleoside-free alpha MEM containing 10% dialyzed fetal calf serum). Approximately 10-15 days post-selection, individual cell clones were picked. Media was analyzed for qp120 expression by radiolabelling the cells with 35S-20 cysteine for 12-18 hours, followed by precipitation of media using a CD4-immunoglobulin-Protein A-Sepharose complex, followed in turn by SDS-PAGE under reducing conditions The levels of gp120 in the media of these (Figure 6). clones were also quantitated (Figure 5) by ELISA performed The method involves coating 96-well plates as follows. overnight with sheep polyclonal IgG against the highly conserved C-terminus of gp120 (D7234, Aalto Bioreagents). After washing, dilutions of a standard gp120 preparation in cell growth medium, or supernatant from the transfected cells, were incubated for 1 hour. The plates were washed again, and incubated for one hour with a horseradish peroxidase-conjugated anti-gp120 monoclonal antibody (9204, DuPont). Following a final wash, the peroxidase substrate OPD (DuPont) was added and the amount

38

of gp120 determined by comparing absorbance of unknowns with a standard curve. Standards were prepared from purified gp120 made in CHO cells, a small quantity of which was obtained from Celltech Ltd. Clones expressing the highest levels were subjected to successive rounds of amplification of the newly introduced DNA sequences in increasing concentrations of methotrexate. Stable CHO cell lines were thus generated which secrete at least 1 microgram/milliliter of HIV-1<sub>1AI</sub> gp120.

10

### 3. Construction of PPI4-tPA-qp120<sub>rg-FL</sub>

a. The HIV-1<sub>LAI</sub> gp120 env nucleotide sequence in PPI4-tpA-gp120<sub>LAI</sub> was replaced by the nucleotide sequence encoding the mature gp120<sub>IR-FL</sub> protein. Using the polymerase chain reaction, the JR-FL sequences were amplified from pUC112-1 (27) using primer 5 (GATCGGCGCCAGAGTAGAAAGTTGTGGGTCAC) and primer 4. The PCR fragment was digested with the restriction endonucleases Nar I and Not I, and the fragment subcloned in between the Nar I and Not I sites in PPI4-tpA-gp120<sub>IR-FL</sub> (Figure 7).

### b. <u>Transient expression</u>.

CosM5 cells grown in DMEM containing 10% fetal calf serum were split to 75% confluence. On the following day, the cells were transfected for 16-20 hours with 10 micrograms of CsCl-purified PPI4-tPA-gp120<sub>JR-PL</sub> DNA by the standard CaPO<sub>4</sub> (5) precipitation technique. After transfection, fresh medium was added to the cells. Analysis of the products synthesized 96-120 hours post-transfection was performed by radiolabelling the transfectants with 35S-cysteine for 12-18 hours, followed by precipitation of media using a CD4-immunoglobulin-Protein A-Sepharose complex, followed by SDS-PAGE under reducing conditions (Figure 4).

## 4. Construction of PPI4-tPA-gp120, V3(1).

The V3 loop in tPA-gp120<sub>LAI</sub> consists of amino acids Cys<sub>306</sub> In the V3<sup>(-)</sup> mutant, the amino acids in through Cys333. between these cysteines are replaced by the pentapeptide 5 sequence Thr-Gly-Ala-Gly-His. Using the Transformer Site-Directed Mutagenesis Kit (Clonetech), the V3 loop sequence in PPI4-tPA-gp120<sub>1A1</sub> is altered using the mutagenic primer 6 (CTGTAGAAATTAATTGTACAGGTGCTGGACATTGTAACATTAGTAGAGC) primer 7 (CTCGAGCATGCATTCGAAGCTCGCTGATC) as a selection Primer 7 changes a unique Xba I site in the 10 backbone of the parent PPI4 plasmid into a unique BstB I Briefly, the mutagenesis method requires incubating of the parent plasmid with the mutagenic primer and the selection primer, denaturing at 100°C for 3 minutes and then 15 chilling on ice. In the presence of buffered deoxynucleotide triphosphates and T4 DNA polymerase, the primers are allowed to initiate the polymerization of one strand of plasmid DNA. T4 DNA ligase is used to seal the newly synthesized DNA strand to form a covalently closed circle. 20 Hybrid plasmids are then transformed into a MutS strain of E. coli that is deficient in mismatch repair. allowing for the growth of transformed cells, purified from the cells and digested with the selection restriction endonuclease, in this case Xba I. 25 plasmids are cleaved by Xba I while the mutant plasmid remains resistant to cleavage by virtue of the Xba I to BstB Digested DNA is then used to transform E. I conversion. coli, and colonies harboring the mutant plasmid are picked. Multiple mutagenic primers can be used in a single round of The amino acid sequence of the modified 30 mutagenesis. protein is shown in Figure 8.

# 5. Construction of PPI4-tPA-gp120<sub>R-FL</sub>-V3(-).

The V3 loop in tPA-gpl20 $_{\rm JR-FL}$  consists of amino acids Cys $_{293}$ 

25

through Cys<sub>377</sub>. In the V3<sup>(\*)</sup> mutant, the amino acids in between these cysteines are replaced by the pentapeptide sequence Thr-Gly-Ala-Gly-His. Using the Transformer Site-Directed Mutagenesis Kit (Clonetech), the V3 loop sequence in PPI4-tPA-gp120<sub>JR-FL</sub> is altered using the mutagenic primer 6 (CTGTAGAAATTAATTGTACAGGTGCTGGACATTGTAACATTAGTAGAGC) and primer 7 as a selection primer. The amino acid sequence of the modified protein is shown in Figure 9.

# 10 6. Construction of PPI4-tPA-qp120, A-CD4(-).

Using the Transformer Site-Directed Mutagenesis Kit (Clonetech), the selection primer 7. and the mutagenic primer 8 (CAATTTATAAACATGGTGCAGGAAGTAGG), Trp<sub>437</sub> of tpA-gp120<sub>LAI</sub>, which is in an equivalent position to the tryptophan residue in the HXBc2 strain of HIV-1, is mutated to a Val in the expression vector PPI4-tpA-gp120<sub>LAI</sub> to generate PPI4-tpA-gp120<sub>LAI</sub>-CD4<sup>(-)</sup>. The sequence for gp120<sub>LAI</sub>-CD4<sup>(-)</sup> is shown in Figure 12.

# 20 7. Construction of PPI4-tPA-qp120<sub>TP,FI</sub>-CD4<sup>(4)</sup>.

In a fashion similar to that described above,  $Trp_{424}$  of  $tPA-gp120_{IR-FL}$  is mutated to a Val in the expression vector  $PP14-tPA-gp120_{IR-FL}$  using the selection primer 7 and the mutagenic primer 9 (CAAATTATAAACATGGTGCAGGAAGTAGG) to generate  $PP14-tPA-gp120_{IR-FL}-CD4^{(-)}$ . The sequence for  $gp120_{IR-FL}-CD4^{(-)}$  is shown in Figure 13.

# 8. Construction of PPI4-tPA-qp120, N-V3()-CD4().

The tPA-gp120<sub>LAI</sub> double mutant, V3<sup>(4)</sup>-CD4<sup>(4)</sup>, is constructed by including the mutagenic primers 6 and 8, and the selection primer 7 simultaneously in the reaction tube with PPI4-tPA-gp120<sub>LAI</sub> as the DNA template. The final construct is named PPI4-tPA-gp120<sub>LAI</sub>-V3<sup>(4)</sup>-CD4<sup>(4)</sup>, and its sequence is shown in figure 10.

41

# 9. Construction of PPI4-tPA-qp120<sub>JR-FL</sub>-V3<sup>(-)</sup>-CD4<sup>(-)</sup>.

The tPA-gp120<sub>JR-FL</sub> double mutant, V3<sup>(·)</sup>-CD4<sup>(·)</sup>, is constructed by including the mutagenic primers 6 and 9, and the selection primer 7 simultaneously in the reaction tube with PPI4-tpA-gp120<sub>JR-FL</sub> as the DNA template. The final construct is named PPI4-tPA-gp120<sub>JR-FL</sub>-V3<sup>(·)</sup>-CD4<sup>(·)</sup>, and its sequence is shown in figure 11.

# 10. Expression of mutant HIV-1 gp120 in mammalian cells.

### 10 a. <u>Transient expression</u>.

CosM5 cells grown in DMEM containing 10% fetal calf serum are split to 75% confluence. On the next day, the cells are transfected for 16-20 hours with 10 micrograms of CsClpurified mutant HIV-1 DNA by the standard CaPO. precipitation technique. After transfection, fresh medium 15 is added to the cells. Analysis of the products synthesized 96-120 hours post-transfection is performed radiolabelling the transfectants with 35-cysteine for 12-18 hours, followed by precipitation of media using a sheep 20 polyclonal IgG against the highly conserved C-terminus of gp120.

# b. Stable expression.

Dhfr Chinese hamster ovary cells (CHO) are transfected with 25 20 micrograms of CsCl-purified DNA encoding the native or mutant HIV-1 gp120 glycoproteins. Approximately 3-5 days post-transfection, cells are placed in selective medium (nucleoside-free alpha MEM containing 10% dialyzed fetal calf serum). Approximately 10-15 days post-selection, individual cell clones are picked. Media is analyzed for gp120 expression by radiolabelling the cells with 35s-cysteine for 12-18 hours, followed by quantitative immunoprecipitation of media using a sheep polyclonal IgG against the highly conserved C-terminus of gp120, followed

by SDS-PAGE under conditions. in reducing Alternatively, one can quantitate the level of gp120 by ELISA performed as follows. The method involves coating 96well plates overnight with sheep polyclonal IgG against the 5 highly conserved C-terminus of gp120 (D7234, Aalto Bioreagents). After washing, dilutions of a standard qp120 preparation in cell growth medium, or supernatant from the stably-transfected cells, are incubated for 1 hour. plates are washed again, and incubated for one hour with a 10 human MoAb (F105, AIDS Research & Reference Reagent Program, No. 857). The plates are washed again, and incubated again for 1 hour with a horseradish-peroxidase-conjugated goat anti-human IgG (Cappel). Following a final wash, the peroxidase substrate OPD (DuPont) is added and the amount of gp120 determined by comparing absorbance of unknowns with a 15 standard curve. Standards are prepared from purified gp120 made in CHO cells, a small quantity of which is obtained from Celltech Ltd. Clones expressing the highest levels are subjected to successive rounds of amplification of the newly introduced DNA sequences in increasing concentrations of Stable CHO cell lines are thus generated methotrexate. which secrete at least 1 microgram/milliliter of mutant HIV-1 gp120.

#### 25 11. Purification of HIV-1 qp120 proteins.

20

A one-step immunoaffinity procedure is used to purify the recombinant gp120 molecules described. Briefly, culture supernatant is collected and clarified by centrifugation. An immunoaffinity column consisting of a matrix coupled to a sheep polyclonal anti-gp120 IgG (D7234, Aalto Bioreagents) 30 directed against the highly conserved C-terminal (APTKAKRRVVQREKR) of gp120 is used to specifically adsorb gp120 from the cell culture media. This antisera recognizes native gp120, the V3 loop deletion mutants, and the CD4(4)

43

mutants since the C-terminal ends of these molecules remain The bound gp120 is then eluted with 2M MgCl, unaltered. concentrated by Amicon filtration, and dialyzed into 10 mM HEPES, pH 7.0. The purity of the proteins is determined by 5 SDS-PAGE and silver staining.

Characterization of recombinant HIV-1 qp120 proteins. The purified glycoproteins are subjected to extensive biochemical and immunologic characterization. The integrity 10 of the proteins is monitored by SDS-PAGE and silver staining reducing and non-reducing conditions. glycoproteins are deglycosylated by treatment with the enzyme N-glycosidase F which cleaves N-linked saccharides, and are assayed by SDS-PAGE and silver staining to monitor molecular weight shifts. The purified glycoproteins are also tested for reactivity with several well characterized anti-gp120 monoclonal antibodies that recognize both linear and discontinuous epitopes. binding affinity to sCD4 is estimated using an ELISA assay.

20

15

The purified proteins HIV-1 gp120<sub>LAI</sub>, gp120<sub>LAI</sub>-V3<sup>(-)</sup>, gp120<sub>LAI</sub>-V3<sup>(-)</sup>  $^{\circ}$ -CD4 $^{\circ}$ , gp120<sub>IR-FL</sub>, gp120<sub>IR-FL</sub>-V3 $^{\circ}$ , and gp120<sub>IR-FL</sub>-V3 $^{\circ}$ -CD4 $^{\circ}$ , were tested for their ability to bind cell surface human CD4. DG44 #3 cells, a recombinant cell line designed to express human CD4 on the membrane surface, were grown in T flasks and trypsinized. 5 X 10<sup>5</sup> cells/experiment were aliquoted into FACS buffer (PBS + 2% BSA and 0.1% NaN3), washed several times in the same buffer, and then incubated with 100 ul of a solution of purified gp120 protein at 5ug/ml in 30 FACS buffer at 37°C for 2 hr. The cells were washed in FACS buffer, and then incubated in 100 ul solution containing 5ug/ml sheep polyclonal IgG against the highly conserved Cterminus of gp120 in FACS buffer at 37°C for 2 hr. cells were washed in FACS buffer then incubated in 100 ul

solution containing FITC-labeled rabbit anti-sheep IgG polyclonal antibody at 37°C for 2 hr. The cells were washed with FACS buffer and then resuspended in 500 ul FACS buffer. The cells were then analyzed on a Becton Dickinson FACScan according to the manufacturer's instructions. As a control for expression of CD4 on the DG44 #3 cells, FITC-labeled OKT4A (Becton Dickinson) was used.

# 13. A protocol for inoculation of animals with the mutant HIV-1 gp120 envelope glycoproteins.

Alum is used as an adjuvant during the inoculation series. The inoculum is prepared by dissolving the mutant HIV-1 gp120 envelope glycoprotein antigen in physiologic saline at a final antigen concentration of 100 ug/ml. Preformed alum (aluminum hydroxide gel) is added to the solution to a final level of 500 ug/ml aluminum. The antigen is allowed to adsorb onto the alum gel for two hours at room temperature. Following adsorption, the gel with the antigen is washed twice with physiologic saline and resuspended in the saline to a protein concentration of 100 ug/ml.

Monkeys and/or Guinea Pigs are individually inoculated with four 100 ug doses of the mutant HIV-1 gp120 envelope glycoprotein antigen adsorbed onto alum. Each dose is injected intramuscularly. The doses are delivered one or five months apart (week 0, 4, 8 and 28). the animals are bled at intervals of two or four weeks. Serum samples are prepared from each bleed to assay for the development of specific antibodies as described in the subsequent sections.

30

35

25

WO 94/22477

10

# 14. Analysis of sera for anti-mutant HIV-1 gp120 envelope glycoprotein IqG antibodies.

Each serum sample is analyzed by ELISA. Polystyrene microtiter plates are coated with 0.5 ug per well of pure mutant HIV-1 gp120 envelope glycoprotein in phosphate-

10

buffered physiological saline (PBS) at 4°C. Each well is then washed with PBS containing 0.5% TWEEN-20 (PBS-TW). Test serum, diluted serially in PBS-TW, is added to the mutant HIV-1 gp120 envelope glycoprotein-containing wells 5 and allowed to react with the adsorbed mutant HIV-1.qp120 envelope glycoprotein for one hour at 37°C. The wells are then washed extensively in PBS-TW. Each well then receives 0.1% p-nitrophenyl phosphate in 10% diethanolamine, pH 9.8, containing 0.5 mM MgCl<sub>2</sub>.6H<sub>2</sub>0. The ensuing reaction is allowed to proceed at room temperature for 30 minutes, at which time it is terminated by the addition of 3.0 N NaOH.

The greater the interaction of antibodies in the test serum with the mutant HIV-1 gp120 envelope glycoprotein, the greater is the amount of alkaline phosphatase bound onto the The phosphatase enzyme mediates the breakdown of pnitrophenyl phosphate into a molecular substance which absorbs light at a wavelength of 405 nm. Hence, there exists a direct relationship between the absorbance at 405 20 nm of light at the end of the ELISA reaction and the amount of mutant HIV-1 gp120 envelope glycoprotein-bound antibody. All animals inoculated with mutant HIV-1 gp120 envelope glycoprotein whose serum reacts specifically with the mutant HIV-1 gp120 envelope glycoprotein in the ELISA have a 25 positive antibody response against mutant HIV-1 gp120 envelope glycoprotein.

### Analysis of sera for activity which specifically 15. neutralizes HIV-1 infectivity.

30 Virus-neutralizing activity is determined with an assay based on the use of multiplicity curves in which the ratio of infectious virus surviving antibody treatment  $(V_n)$  is compared to infectious virus in uninhibited cultures (Va) at various dilutions of antisera. The neutralization titer of

the sera is then interpolated as that sera dilution which yields one log reduction in infectious titer (i.e.,  $V_p/V_q =$ Briefly, 4-fold dilutions of virus (laboratoryadapted and primary isolates) are prepared to yield 5 infectious doses of 0.1 to 100 TCID<sub>50</sub> (Tissue Culture Infection Dose) in 20 ul. Serial 3-fold dilutions of sera are also prepared and 20 ul of each serum dilution are incubated with each dilution of virus in duplicate for 60 minutes at room temperature in a 96-well microtiter plate. 10 20 ul of AA5 cells (PHA stimulated PBMCs for primary HIV-1 isolates) are then added to the serum/virus mixtures. Cells are cultured for 7 days by the addition of fresh medium every other day. On the seventh day, supernatant from each well is removed and tested for the presence of reverse transcriptase (RT). Infection in each well is then scored 15 as either positive or negative based on the RT counts, and the infectious dose of virus in each treatment group is calculated using the Reed and Muench (28) formula. The neutralization titers represent the reciprocal serum 20 dilution required to reduced infectious dose of virus by one log. The above culture time is for the prototypic HIV-1, A isolate tested on the AA5 cell line. In the case of primary isolates, the termination date is usually 11-14 days. Culture conditions for PBMCs is not as demanding since 25 doubling time is restricted. In the case of PBMCs, one day PHA stimulations are used at a final concentration of 1.5 X 106/ml on day 0. Half that number of fresh PBMCs are then added again on days 4 and 8. This multiple addition of PBMCs is meant to amplify virus output upon successful 30 infection so that the readout RT signal is strong. the final readout titer for the primary isolate/PBMC is the reciprocal serum dilution which reduces infectious titer by one log.

### 16. Passive hyperimmune therapy.

Non-HIV-1-infected humans are immunized with the mutant HIV1 gp120 envelope glycoprotein antigens according to a
protocol similar to that described above in section 12. For
5 passive hyperimmune therapy in HIV-1-infected individuals,
blood plasma is taken from mutant HIV-1 gp120 envelope
glycoprotein immunized, non-HIV-1-infected human donors
whose plasma has high levels of neutralizing antibodies.
The plasma is pooled from several donors, purified to remove
10 nonimmunoglobulin proteins and is then sterilized to kill
any other viruses or pathogens. The treated plasma is then
injected into individuals infected with HIV-1, with repeated
injections every week, every two weeks, or every month.

48

### Results

Eukaryotic expression vectors designed to express high levels of HIV-1<sub>LAI</sub> gp120 and HIV-1<sub>JR-FL</sub> gp120 were constructed. The CMV MIE promoter/enhancer was used to drive the transcription of a gene fusion consisting of the human tPA signal sequence fused to mature gp120 (Figures 2 and 7). The complete sequence of the transcription unit from the Hinc II site of the CMV promoter/enhancer to the Not I site 10 just 3' from the stop codon in gp120 is shown in figure 3. This vector was used to transfect COSM5 cells in a transient assay. The transfected cells were labeled with 35S-cysteine and the media immunoprecipitated with a CD4-immunoglobulin-Protein A-Sepharose complex. The precipitated products were 15 analyzed using a reducing 10% SDS-PAGE autoradiography (Figure 4). A 120 kD band was detected when PPI4-tPA-gp120<sub>LAI</sub> was used to transfect COS cells (lane 3). A band migrating with a slightly lower molecular mass was detected when PPI4-tPA-gp120 $_{\mbox{\scriptsize IR-FL}}$  was used to transfect COS 20 cells (lane 4). No radiolabeled products were detected in the mock infected cells. Using a sheep polyclonal antibody directed against the highly conserved C-terminal end of HIV-1 gp120 in an ELISA assay, the level of expression of HIV-1 qp120 was determined to be 2350 ng/ml.

25

30

The PPI4-tPA-gp120<sub>LAI</sub> vector was then used to stably transfect the dhfr CHO cell line DXB11. Two days post-transfection, the cells were plated at low density in nucleoside-free medium. Eight days post-transfection, surviving clones were isolated and expanded. Individual primary transfectants were tested for gp120 expression using the ELISA method described in the methods section. Several primary CHO transfectants expressed significant quantities (10-120 ng/ml) of gp120 (Figure 5). Three of the highest

49

expressing clones were then subjected to increasing concentrations of methotrexate in order to amplify, in tandem, the copy number of the dhfr and gpl20 genes. Cell lines were established that express high levels of gpl20 with rates of secretion greater than 1 mg/liter. These were then used to purify gpl20 to homogeneity.

Six CHO cell lines were established, using the procedures described in the methods sections, that express high levels of the following proteins: HIV-1 gp120<sub>LAI</sub>, gp120<sub>LAI</sub>-V3<sup>(·)</sup>, gp120<sub>LAI</sub>-V3<sup>(·)</sup>, gp120<sub>JR-FL</sub>, gp120<sub>JR-FL</sub>-V3<sup>(·)</sup>, and gp120<sub>JR-FL</sub>-V3<sup>(·)</sup>. CD4<sup>(·)</sup>. Metabolic labeling of these cells with <sup>35</sup>S-cysteine followed by immunoprecipitation with the human monoclonal antibody F105 and analyzed by SDS-PAGE and autoradiography showed the presence of the gp120 proteins in the culture supernatant (Figure 14). From these cell lines the gp120 proteins were purified to homogeneity. Analysis by SDS-PAGE followed by silver-staining showed the purity of these proteins to be greater than 90% (Figure 15).

20

25

It was shown by FACScan analysis that the two CD4 binding mutants  $HIV-1gp120_{IAI}-V3^{(\cdot)}-CD4^{(\cdot)}$  and  $HIV-1gp120_{IR-FL}-V3^{(\cdot)}-CD4^{(\cdot)}$  had no appreciable binding to recombinant cell lines designed to express high levels of human CD4 on their membrane surface (Figure 16, panel 4 and data not shown, respectively).

### Discussion

The advantage of using the mutant HIV-1 gp120 envelope glycoproteins as immunogens is that these proteins will not elicit an immune response against the V3 loop, a highly immunodominant epitope on gp120. This is significant because the V3 loop may skew the humoral immune response away from discontinuous epitopes in the CD4-binding site. Mutant HIV-1 10 gp120 envelope glycoproteins having partial and total v3 loop deletions have been made (30). Deletion of the V3 loop therefore exposes the CD4-binding site to the immune system. allowing the immune system to mount a response against this critical region (18). Another advantage of using the mutant HIV-1 gp120 envelope glycoprotein as an immunogen is that it 15 has significantly reduced affinity for cell surface CD4. An efficient humoral immune response depends on the binding of antigen to B cell surface immunoglobulin. The presence of the high-affinity CD4 receptor on large numbers of cells in 20 the body may significantly diminish the ability of native gp120 to induce an effective humoral immune response. rationale of mutating gp120 at the CD4 binding site is to redirect the mutant HIV-1 gp120 envelope glycoprotein away from cell surface CD4 toward immunoglobulin-bearing B cells, thereby allowing the immune system to mount a response against, inter alia, the CD4-binding site.

### References

1. Klatzmann, D.R., et. al. (1990) Immunodeficiency Reviews 2, 43-66.

5

- 2. Lasky, L.A., et. al. (1987) Cell 50, 975-985.
- 3. Maddon, P.J., et. al. (1986) Cell 47, 333-348.
- 10 4. Maddon, P.J., et. al. (1988) Cell 54, 865-874.
  - 5. Maddon, P.J., et. al. (1985) Cell 42, 93-104.
- 6. Maddon, P.J., et. al. (1987) Proc. Natl. Acad. Sci. U.S.A. 84, 9155-9159.
  - 7. Richardson, N.E., et. al. (1988) Proc. Natl. Acad. Sci. U.S.A. 85, 6102-6106.
- 20 8. Chao, B.H., et. al. (1989) J. Biol. Chem. 264, 5812-5817.
  - 9. Arthos, J., et. al. (1989) Cell 57, 469-481.
- 25 10. Wang, J., et. al. (1990) Nature 348, 411-418.
  - 11. Ryu, S.-E., et. al. (1990) Nature 348, 419-426.
- 12. Leonard, C.K., et. al. (1990) J. Biol. Chem. 265, 30 10373-10382.
  - 13. Earl, P.L., et. al. (1990) Proc. Natl. Acad. Sci. U.S.A. 87, 648-652.
- 35 14. Helseth, E., et. al. (1991) J. Virol. 65, 2119-2123.

52

- 15. Bolognesi, D.P. (1990) TIBTech 8, 40-45.
- 16. Olshevsky, U., et. al. (1990) J. Virol. 64, 5701-5707.
- 5 17. Steimer, K.S., et. al. (1991) AIDS 5, S135-143.
  - 18. Wyatt, R., et. al. (1992) J. Virol. 66, 6997-7004.
- Zolla-Pazner, S., et. al. (1992) Sem. in Virology 3,
   203-211.
  - 20. Steimer, K.S., et. al. (1991) Science 254, 105-108.
  - 21. Pollard, S.R., et. al. (1992) EMBO J. 11, 585-591.
- 22. Okayama, H. (1983) Mol. Cell. Biol. 3, 280-289.

15

20

25

- 23. Maniatis, T., et. al. (1990) Molecular Cloning, Vol. 1-3.
- 24. Thomsen, D.R., et. al. (1984) Proc. Natl. Acad. Sci. U.S.A. 81, 659-663.
  - 25. Pennica, D., et. al. (1983) Nature 301, 214-221.
- 26. Wain-Hobson, S., et. al. (1985) Cell 40, 9-17.
  - 27. Koyanagi, Y. (1987) Science 236, 819-822.
- 30 28. Reed, L.J. (1938) Am. J. Hyg., 27, 493-497.
  - 29. Cohn, E.J. et al., (1944) J. Clin. Invest. 23, 417-432.
- 30. Shiow-Her, C., et al. (1992) J. of Cellular Biochem., 35 Supplement 16E, Abstrtact Q105.

### SEQUENCE LISTING

- (1) GENERAL INFORMATION:
  - (i) APPLICANT: Progenics Pharmaceuticals, Inc.
  - (ii) TITLE OF INVENTION: HIV-1 VACCINES, ANTIBODY COMPOSITIONS RELATED THERETO, AND THERAPEUTIC AND PROPHYLACTIC USES THEREOF
  - (iii) NUMBER OF SEQUENCES: 29
  - (iv) CORRESPONDENCE ADDRESS:
    - (A) ADDRESSEE: Cooper & Dumham
    - (B) STREET: 30 Rockefeller Plaza
    - (C) CITY: New York
    - (D) STATE: New York
    - (E) COUNTRY: USA
    - (F) ZIP: 10112
  - (v) COMPUTER READABLE FORM:
    - (A) MEDIUM TYPE: Floppy disk
    - (B) COMPUTER: IBM PC compatible
    - (C) OPERATING SYSTEM: PC-DOS/MS-DOS
    - (D) SOFTWARE: Patentin Release #1.24
  - (vi) CURRENT APPLICATION DATA:
    - (A) APPLICATION NUMBER:
      - (B) FILING DATE:
      - (C) CLASSIFICATION:
  - (vii) PRIOR APPLICATION DATA:
    - (A) APPLICATION NUMBER: US 08/037,816
    - (B) FILING DATE: 26-MAR-1993
- (viii) ATTORNEY/AGENT INFORMATION:

  - (A) NAME: White, John P. (B) REGISTRATION NUMBER: 28,678
  - (C) REFERENCE/DOCKET NUMBER: 41190-A-PCT/JPW/AJM
  - (ix) TELECOMPUNICATION INFORMATION:
    - (A) TELEPHONE: (212) 977-9550
    - (B) TELEFAX: (212) 664-0525
    - (C) TELEX: 422523 COOPUL
- (2) INFORMATION FOR SEQ ID NO:1:
  - (i) SEQUENCE CHARACTERISTICS:
    - (A) LENGTH: 45 mains acids
    - (B) TYPE: amino acid
    - (C) STRANDEDNESS: single (D) TOPOLOGY: linear
  - (ii) MOLECULE TYPE: protein
  - (xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:
  - X12 Xaa Xaa Cys Xaa Ile Xaa Xaa Xaa Xaa Xaa Xaa Trp Xaa Xaa Xaa
  - Xaa Xaa Ala Xaa Tyr Xaa Xaa Pro Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
  - Ser Xaa Xaa Thr Gly Xaa Xaa Xaa Arg Xaa Gly Xaa

54

35

45

- (2) INFORMATION FOR SEQ ID NO:2:
  - (i) SEQUENCE CHARACTERISTICS:
    - (A) LENGTH: 45 amino acids
      (B) TYPE: amino acid
    - (C) STRANDEDNESS: single
    - (D) TOPOLOGY: linear
  - (ii) MOLECULE TYPE: protein
  - (xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Thr Leu Pro Cys Arg Ile Lys Gln Phe Ile Asn Met Trp Gln Glu Val

40

Gly Lys Ala Met Tyr Ala Pro Pro Ile Ser Gly Gln Ile Arg Cys Ser 20 25 30

Ser Asn lie Thr Gty Leu Leu Leu Thr Arg Asp Gty Gly 35 40

- (2) INFORMATION FOR SEQ ID NO:3:
  - (i) SEQUENCE CHARACTERISTICS:
    - (A) LENGTH: 45 amino acids
    - (B) TYPE: amino acid
    - (C) STRANDEDNESS: single
    - (D) TOPOLOGY: linear
  - (ii) MOLECULE TYPE: protein
  - (xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

Thr Leu Pro Cys Arg Ile Lys Gln Ile Ile Asn Met Trp Gln Glu Val 1 5 10 15

Gly Lys Ala Met Tyr Ala Pro Pro Ile Arg Gly Gln Ile Arg Cys Ser 20 25 30

Ser Asn Ile Thr Gly Leu Leu Leu Thr Arg Asp Gly Gly 35 40 45

- (2) INFORMATION FOR SEQ ID NO:4:
  - (i) SEQUENCE CHARACTERISTICS:
    - (A) LENGTH: 45 base pairs
    - (8) TYPE: nucleic acid (C) STRANDEDNESS: single
    - (D) TOPOLOGY: linear
  - (ii) MOLECULE TYPE: DNA (genomic)
  - (xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

CATCCTGCAG TCACCGTCCT TGACACGATG GATGCAATGA AGAGA

(2) INFORMATION FOR SEQ ID NO:5:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 36 base pairs
  - (B) TYPE: nucleic acid

(C) STRANDEDNESS: single (D) TOPOLOGY: linear	
(ii) MOLECULE TYPE: DNA (genomic)	
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:	
AAGTETTETE ETEGGTETTG TETTTTTAAC ACCEAG	36
(2) INFORMATION FOR SEQ ID NO:6:	
(i) SEQUENCE CHARACTERISTICS:  (A) LENGTH: 36 base pairs  (B) TYPE: nucleic acid  (C) STRANDEDNESS: single  (D) TOPOLOGY: linear	
(ii) MOLECULE TYPE: DNA (genomic)	
(xf) SEQUENCE DESCRIPTION: SEQ ID NO:6:	
TTCAGAAGAG GAGCCAGAAC AGAAAAATTG TGGGTC	36
(2) INFORMATION FOR SEQ ID NO:7:	
(i) SEQUENCE CHARACTERISTICS:  (A) LENGTH: 39 base pairs  (B) TYPE: nucleic acid  (C) STRANDEDNESS: single  (D) TOPOLOGY: linear	
(ii) MOLECULE TYPE: DNA (genomic)	
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:	
GGAAAAAAGC GGCCGCTCAT TTTTCTCTCT GCACCACTC	39
(2) INFORMATION FOR SEQ ID NO:8:	
(i) SEQUENCE CHARACTERISTICS:  (A) LENGTH: 33 base pairs  (B) TYPE: nucleic acid  (C) STRANDEDNESS: single  (D) TOPOLOGY: linear	·
(ii) MOLECULE TYPE: DNA (genomic)	
(xi) SEQUENCE DESCRIPTION: SEQ 1D NO:8:	
GATCGGCGCC AGAGTAGAAA AGTTGTGGGT CAC	33
(2) INFORMATION FOR SEG ID NO:9:	
(i) SEQUENCE CHARACTERISTICS:  (A) LENGTH: 49 base pairs  (B) TYPE: nucleic acid  (C) STRANDEDNESS: single  (D) TOPOLOGY: linear	

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEO ID NO:9:	
CTGTAGAAAT TAATTGTACA GGTGCTGGAC ATTGTAACAT TAGTAGAGC	49
(2) INFORMATION FOR SEQ ID NO:10:	
(i) SEQUENCE CHARACTERISTICS:  (A) LENGTH: 29 base pairs  (B) TYPE: nucleic acid  (C) STRANDEDNESS: single  (D) TOPOLOGY: linear	
(ii) MOLECULE TYPE: DNA (genomic)	
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:	
CTCGAGCATG CATTCGAAGC TCGCTGATC	29
(2) INFORMATION FOR SEQ ID NO:11:	
(i) SEQUENCE CHARACTERISTICS:  (A) LENGTH: 29 base pairs  (B) TYPE: nucleic acid  (C) STRANDEDNESS: single  (D) TOPOLOGY: linear	
(ii) NOLECULE TYPE: DNA (genomic)	
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:	
CAATTTATAA ACATGGTGCA GGAAGTAGG	\ 29
(2) INFORMATION FOR SEG ID NO:12:	
(i) SEQUENCE CHARACTERISTICS:  (A) LENGTH: 29 base pairs  (B) TYPE: nucleic acid  (C) STRANDEDNESS: single  (D) TOPOLOGY: linear	
(ii) MOLECULE TYPE: DMA (genomic)	
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:	
CAAATTATAA ACATGGTGCA GGAAGTAGG	29
(2) INFORMATION FOR SEG ID NO:13:	
(i) SEGUENCE CHARACTERISTICS: (A) LENGTH: 3125 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear	
(ii) MOLECULE TYPE: DNA (genomic)	
(ix) FEATURE: (A) NAME/KEY: CDS (B) LOCATION: 15553115 (D) OTHER INFORMATION:	

### (xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

TTGACATTGA TTATTGACTA GTTATTAATA GTAATCAATT ACGGGGTCAT TAGTTCATAG	60
CCCATATATG GAGTTCCGCG TTACATAACT TACGGTAAAT GGCCCGCCTG GCTGACCGCC	120
CAACGACCCC CGCCCATTGA CGTCAATAAT GACGTATGTT CCCATAGTAA CGCCAATAGG	120
GACTITCCAT IGACGICAAT GGGIGGACTA TITACGGIAA ACTGCCCACT IGGCAGTACA	240
TCAAGTGTAT CATATGECAA GTACGCCCCC TATTGACGTC AATGACGGTA AATGGCCCGC	300
CTGGCATTAT GCCCAGTACA TGACCTTATG GGACTTTCCT ACTTGGCAGT ACATCTACGT	360
ATTAGTCATC GCTATTACCA IGGTGATGCG GTTTTGGCAG TACATCAATG GGCGTGGATA	420
GCGGTTTGAC TCACGGGGAT TTCCAAGTCT CCACCCCATT GACGTCAATG GGAGTTTGTT	480
TTGGCACCAA AATCAACGGG ACTITECAAA ATGTCGTAAC AACTCCGCCC CATTGACGCA	540
AATGGGEGGT AGGEGTGTAC GGTGGGAGGT CTATATAAGC AGAGCTCGTT TAGTGAACCG	600
TCAGATCGCC TGGAGACGCC ATCCACGCTG TTTTGACCTC CATAGAAGAC ACCGGGACCG	660
ATCCAGCETC CGCGGCCGGG AACGGTGCAT TGGAACGCGG ATTCCCCGTG CCAAGAGTGA	720
EGTAAGTACE GESTATAGAS TETATAGGEA CACCESTITG GETETTATGE ATGETATACT	780
GTTTTTGGCT TGGGCCAACA CCCCGTCCTA GATAGGTGAT GGTATAGCTT AGCCTATAGG	840
TGTGGGTTAT TGACCATTAT TGACCACTCC CCTATTGGTG ACGATACTTT CCATTACTAA	900
TCCATAACAT GGCCGCTCTT TGCCACAACT ATCTCTATTG GCTATATGCC AATACTCTGT	960
CCTTCAGAGA CTGACACGGA CTCTGTATIT TTACAGGATG GGGTCCCATT TATTATTTAC	1020
AMATTCACAT ATACAACAAC GCCGTCCCCC GTGCCCGCAG TTTTTATTAA CATGCGGGAT	1080
CTCCACGCGA ATCTCGGGTA CGTGTTCCGG ACATGGGCTC TTCTCCGGTA GCGGCGGAGC	1140
TECACATECG AGESTGTESS ATGSSSCATGS STOCKAGEGGS TEATGGTEGG TEGGCAGSTS	1200
CTTGCTCCTA ACAGTGGAGG CCAGACTTAG GCACAGGACA ATGCCCACCA CCACCAGTGT	1260
GCCGCACAAG GCCGTGGCGG TAGGGTATGT GTCTGAAAAT GAGCTCGGAG ATTGGGCTCG	1320
CACCGCTGAC GCAGATGGAA GACTTAAGGC AGCGGCAGAA GAAGATGCAG GCAGCTGAGT	1380
TGTTGTATTC TGTAGAGTTG GAGGTAACTC CCGTTGCGGT GCTGTTAACG GTGGAGGGCA	1440
GTGTAGTCTG AGCAGTACTC GTTGCTGCCG CGCGCGCCAC CAGACATAAT AGCTGACAGA	1500
CTANCAGACT GTTCCTTTCC ATGGGTCTTT TCTGCAGTCA CCGTCCTTGA CACG ATG Met 1	1557
CAT GCA ATG AAG AGA GGG CTC TGC TGT GTG CTG CTG TGT GGA GCA ASP Ala Het Lys Arg Gly Leu Cys Cys Val Leu Leu Cys Gly Ala 5 10 15	1605
GTC TTC GTT TCG CCC AGC CAG GAA ATC CAT GCC CGA TTC AGA AGA GGC Val Phe Val Ser Pro Ser Gin Glu Ile His Ala Arg Phe Arg Arg Gly 20 25 30	1653
GCC AGA ACA GAA AAA TTG TGG GTC ACA GTC TAT TAT GGG GTA CCT GTG Ala Arg Thr Glu Lys Leu Trp Val Thr Val Tyr Tyr Gly Val Pro Val 35 40 45	1701
TGG AAG GAA GCA ACC ACC ACT CTA TTT TGT GCA TCA GAT GCT AAA GCA	1749

T r		ys (	ilu .	Ala	Th		r 11	ir Le	eu Ph	ne Cy	s Al		r Asp	Ala	Lys	Ala 65	
						LHi				G GC P Al 7	a In					Рго	1797
									l Va	A TT IL Le O					Glu		1845
		n M							t Va	A GA							1893
		r L						r Le		G CC/ s Pro			Lys				1941
	Cy						Cy			T TT(		ASI					1989
				sn						G GGG F Gly 155	Glu					AAA Lys	2037
GGA Gly	GA	G A1	e L	AA i ys i 65	AAC Asn	TGC Cys	TC1 Ser	TTC Phe	AA1 AST 170	T ATO	AGC Ser	ACA Thr	AGC Ser	ATA Ile 175	AGA Arg	GGT Gly	2085
AAG Lys	GT (	G CA L GL 18	n Ly	AA (	GAA Glu	TAT Tyr	GCA Ala	777 Phe 185	Phe	TAT	Lys	CTT Leu	GAT ASP 190	ATA	ATA Ile	CCA Pro	2133
ATA	GA1 ASE 195	) As	T GA	AT #	hr	ACC Thr	AGC Ser 200	Tyr	ACG Thr	TTG Leu	ACA Thr	AGT Ser 205	TGT Cys	AAC Asn	ACC Thr	TCA Ser	2181
Val 210	He	Th	r Gi	n A	la	Cys 215	Pro	Lys	Val	TCC Ser	Phe 220	GLu	Pro	He	Pro	11e 225	2 <b>229</b>
CAT His	TAT	TG'	GC S Al	a P	CG 70 30	GCT Ala	GGT Gly	TTT	GCG Ala	ATT Ile 235	CTA Leu	AAA Lys	TGT Cys	AAT ASN	AAT Asn 240	AAG Lys	2277
ACG Thr	TTC Phe	ASI	GG G G L 24	y I	CA (	GGA Sly	CCA Pro	TGT Cys	ACA Thr 250	AAT Asn	GTC Val	AGC Ser	ACA Thr	GTA Val 255	CAA Gln	TGT Cys	2325
T			11			·				ACT Thr			_				2373
Ser	CTA Leu 275	GCA Ala	GA	A G/ u Gl	M G	lu	GTA Val 280	GTA Vai	ATT	aga Arg	Ser	GCC Ala 285	AAT Asn	TTC Phe	ACA Thr	GAC Asp	2421
AAT Asn 290	Ala	Lys	Thr	- 11	e 1 2	le \ 95	Val	Gln	Leu	Asn	Gl n 300	Ser	Val	Glu	Ile	Asn 305	24 <del>69</del>
TGT / Cys					n A				Arg					Ile			2517
GGA ( Gly F	Pro	GGG Gly	AGA Arg	GC	A T' a Pi	TT G	at '	ACA Thr	ATA ile	GGA Gly	AAA A Lys	ATA Ile	GGA Gly	AAT Asn	ATG Met	AGA Arg	2565

59 ·

			325					330					335			
CAA Gln	GCA Ala	CAT His 340	Cys	AAC	ATT	AGT Ser	AGA Arg 345	GCA Ala	AAA Lys	TGG	AAT Asn	GCC Ala 350	ACT Thr	TTA Leu	AAA Lys	2613
CAG Gln	A:A 1 (e 355	GCT Ala	AGC Ser	KAA Lys	TTA Leu	aga arg 360	GAA Glu	CAA	TTT Phe	GGA Gly	AAT Asn 365	AAT ASN	AAA Lys	ACA Thr	ATA ile	2661
ATC Ile 370	TTT Phe	AAG Lys	CAA Gin	TCC Ser	TCA Ser 375	GGA Gly	GGG Gly	GAC Asp	CCA Pro	GAA Glu 380	ATT Ile	GTA Val	ACG Thr	CAC His	AGT Ser 385	2709
777 Phe	AAT Asn	TGT Cys	GGA Gly	GGG Gly 390	GAA Glu	171 Phe	TTC Phe	TAC Tyr	7G7 Cys 395	AAT Asn	TCA Ser	ACA Thr	CAA Gln	CTG Leu 400	TTT Phe	2757
					AAT Asn											2805
ACT Thr	GÁÁ Glu	GGA Gly 420	AGT Ser	GAC Asp	ACA Thr	ATC Ile	ACA Thr 425	CTC Leu	CCA Pro	TGC Cys	AGA Arg	ATA 1 le 430	AAA Lys	CAA Gln	TTT Phe	2853
					GAA Glu											2901
					TGT Cys 455											2949
					AAC Asn											2 <del>99</del> 7
					GAC Asp											3045
					CCA Pro											3093
Arg					GAA Glu		T GA	GCGG	CCGC	:						3125

# (2) INFORMATION FOR SEG ID NO:14:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 520 amino acids (B) TYPE: amino acid

  - (D) TOPOLOGY: Linear

### (ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

Met Asp Ata Met Lys Arg Gly Leu Cys Cys Val Leu Leu Leu Cys Gly 1 5 10 15

Ala Val Phe Val Ser Pro Ser Gln Glu Ile His Ala Arg Phe Arg Arg 20 25 30

Gly Ala Arg Thr Glu Lys Leu Trp Val Thr Val Tyr Tyr Gly Val Pro

Val Tro Lys Glu Ala Thr Thr Thr Leu Phe Cys Ala Ser Asp Ala Lys 50 55 60 Ala Tyr Asp Thr Glu Val His Asn Val Trp Ala Thr His Ala Cys Val 55 70 75 80 Pro Thr Asp Pro Asn Pro Gln Glu Val Val Leu Val Asn Val Thr Glu 85 90 95 Asn Phe Asn Met Trp Lys Asn Asp Met Val Glu Gln Met His Glu Asp 100 105 110 Ile Ile Ser Leu Trp Asp Gln Ser Leu Lys Pro Cys Val Lys Leu Thr 115 120 125 Pro Leu Cys Val Ser Leu Lys Cys Thr Asp Leu Gly Asn Ala Thr Asn 130 135 140 Thr Asn Ser Ser Asn Thr Asn Ser Ser Ser Gly Glu Met Net Met Glu 145 150 155 160 Lys Gly Glu Ile Lys Asn Cys Ser Phe asn Ile Ser Thr Ser Ile Arg 165 170 175 Gly Lys Val Gln Lys Glu Tyr Ala Phe Phe Tyr. Lys Leu Asp Ile Ile 180 185 190 Pro 11e Asp Asn Asp Thr Thr Ser Tyr Thr Leu Thr Ser Cys Asn Thr 195 200 205 Ser Val Ile Thr Gln Ala Cys Pro Lys Val Ser Phe Glu Pro Ile Pro 210 220 Ile His Tyr Cys Ala Pro Ala Gly Phe Ala Ile Leu Lys Cys Asn Asn 225 230 235 240

Lys Thr Phe Asn Gly Thr Gly Pro Cys Thr Asn Val Ser Thr Val Gln 245 250 255

Cys Thr His Gly Ile Arg Pro Val Val Ser Thr Gln Leu Leu Leu Asn 260 265 270

Gly Ser Leu Ala Glu Glu Glu Val Val Ile Arg Ser Ala Asn Phe Thr 275 280 285

Asp Asn Ala Lys Thr Ile 1le Val Gln Leu Asn Gln Ser Val Glu Ile 290 295 300

Asn Cys Thr Arg Pro Asn Asn Asn Thr Arg Lys Ser Ite Arg Ite Gin 305 310 315 320

Arg Gly Pro Gly Arg Ala Phe Val Thr Ile Gly Lys Ile Gly Asn Met 325 330 335

Arg Gln Ala His Cys Asn Ile Ser Arg Ala Lys Trp Asn Ala Thr Leu 340 345 350

Lys Gin Ile Ala Ser Lys Leu Arg Glu Gin Phe Gly Asn Asn Lys Thr 355 360 365

lle Ile Phe Lys Gln Ser Ser Gly Gly Asp Pro Glu Ile Val Thr His 370 380

Ser Phe Asn Cys Gly Glu Phe Phe Tyr Cys Asn Ser Thr Gln Leu 385 390 395 400

Phe Asn Ser Thr Trp Phe Asn Ser Thr Trp Ser Thr Glu Gly Ser Asn

61 -

				4	05				410	)				415		
As	n Th	ır G		ly So 20	er A	sp 1	hr I	le Th 42		J Pro	Cys	Arg	11e 430		Gln	
Ph	e Il		sn Ho 35	et T	rp G	in G	lu Va 44		y Ly:	s Ala	Met	1yr 445		Pro	Pro	
11	e Se 45		ly G	ln I	le A		ys Se <b>55</b>	er Se	r Ası	n Ile	1hr 460		Leu	Leu	Leu	
Th:		g As	sp Gi	y Gl		sn A 70	SN AS	in Asi	n Gly	/ Ser 475		lle	Phe	Arg	Pro 480	
Gl	y Gl	y G	y As	P Me	_	rg A	sp As	n In	490		Glu	Leu	Туг	Lys 495	Туг	
			50	0			ro Le	50		Ala	Pro	Thr	Lys 510	Ala	Lys	
	-	51	5				lu Ly 52	0							•	
(2)	) IN	FORM	ATIC	N FO	Z SE	0 1	NO:	15:								
	(	i) S	(A) (B) (C)	LENG TYPE STRA	TH: : nu NDED	1532 Iclei NESS	RIST Das c ac	e pai id ngle	irs							
			(D)	טישו	LUGI	2 11	near									
	(ii	i) M	OLEC	ULE	TYPE	: DA	IA (g	enomi	c)							
	(i)	() F	EATU	RE:												
			(4)		/KEY	- 0	8									
				NAME.	TION	: 1.	.152									
			(B)	NAME.	TION	: 1.										
	(xi		(B) (D)	NAME, LOCA OTHE	TION	: 1. FORM	.152	<b>4</b> =	ID N	0: 15	:					
ATG	GAT	; ) S	(B) (D) EQUE	NAME, LOCA DTHEI NCE I	TION R IN DESC G AG	: 1. FORM RIPT A GG	.152; IATIO ION: G CT(	SEQ TGC	TGT	GTG	CTG	CTG	стб	ŤGT	GGA	4
ATG Met	GAT Asp	; ) S	(B) (D) EQUE	NAME, LOCA DTHEI NCE I	TION R IN DESC G AG	: 1. FORM RIPT A GG	.152; ATIO	SEQ TGC	TGT	GTG	CTG	CTG Leu	CTG Leu	TGT Cys 15	GGA Gly	<b>.</b>
Met 1 GCA	GAT ASP	GC.	(B) (D) EQUE A AT( B Me	NAME. LOCATOTHEI  NICE II  G AAG Lys	TION R IN DESC G AG G Ar	: 1. FORM RIPT A GG g Gl	.152; IATION: ION: G CT( y Lec	SEQ TGC Cys	TGT Cys 10	GTG Val	CTG Leu GCC	Leu	Leu TTC	Cys 15 AGA	GLY	. 90
Met 1 GCA	GAT ASP	GC.	(B) (D) EQUE A AT( B Me	NAME. LOCA' OTHER NCE I S AAG L Lys	TION R IN DESC G AG G Ar	: 1. FORM RIPT A GG g Gl	.152; IATION: ION: G CT( y Lec	SEQ TGC Cys	TGT Cys 10 ATC	GTG Val	CTG Leu GCC	Leu	Leu TTC	Cys 15 AGA	GLY	
Met 1 GCA Ala GGC	GAT Asp GTC Val	GC: AL: TTI Pho	(B) (D) EQUE A AT( B Me*	NAME, LOCA: OTHER NCE I G AAG L LYS ST TCE S Ser	TION R IN DESC G AC G Ar G CC	: 1. FORM RIPT A GG G GL C AG C Se	.152: IATION: ION: G CT( y Let C CAC F Glr	SEQ TGC Cys GAA Glu 25	TGT Cys 10 ATC Ile	GTG Val CAT His	CTG Leu GCC Ala	CGA Arg	TTC Phe 30 GGG	Cys 15 AGA Arg	GLY AGA Arg	
Met 1 GCA Ala GGC	GAT Asp GTC Val	GC: AL: TTI Pho	(B) (D) (D) EQUE: A AT( B Me: C GTI E Val	NAME, LOCA: OTHER NCE I G AAG L LYS ST TCE S Ser	TION R IN DESC G AC G Ar G CC	: 1. FORM RIPT A GG G GL C AG C Se	.152: IATION: ION: G CTU Y Leu C CAC	SEQ TGC Cys GAAA Glu 25 GTC	TGT Cys 10 ATC Ile	GTG Val CAT His	CTG Leu GCC Ala	CGA Arg	TTC Phe 30 GGG	Cys 15 AGA Arg	GLY AGA Arg	9(
Met 1 GCA Ala GGC Gly	GAT Asp GTC Val GGC GLy	GC. AL. TTT Pho	EQUE: A ATO B Me C GTI C Val C Val C Val C GTI C Val C	NAME LOCATION OF THE STATE OF T	TION R IN	: 1. FORM RIPT A GG G GL C AG S Se S TTI	.152; IATION: ION: G CT( y Let C CA( r Glr G TGG U Trp 40	SEQ CYS GAAA GLU 25 GCTC Val	TGT Cys 10 ATC Ile ACA Thr	GTG Val CAT His GTC Val	CTG Leu GCC Ala TAT Tyr	CGA Arg TAT Tyr 45	TTC Phe 30 GGG GLY	Cys 15 AGA Arg GTA Val	GLY AGA Arg CCT Pro	9(
Met 1 GCA Ala GGC Gly	GAT Asp GTC Val GGC GLy	GC. AL. TTI Pho	EQUE: A ATO B Me C GTI C Val C Val C Val C GTI C Val C	NAME LOCATION OF THE STATE OF T	TION R IN	: 1. FORM RIPT A GG G GL C AG S Se S TTI	.152: ION: ION: G CTG y Lec C CAG F Glr G TGG 40 C ACT F The	SEQ CYS GAAA GLU 25 GCTC Val	TGT Cys 10 ATC Ile ACA Thr	GTG Val CAT His GTC Val	CTG Leu GCC Ala TAT Tyr	CGA Arg TAT Tyr 45	TTC Phe 30 GGG GLY	Cys 15 AGA Arg GTA Val	GLY AGA Arg CCT Pro	94
GCA Ala GGC Gly GTG Val	GAT Asp GTC Val GGC GLy TGG Trp 50	AGAT	(B) (CD) (CD) (CD) (CD) (CD) (CD) (CD) (CD	NAME, LOCATOR	TION TION TO THE TENT TO THE TENT TENT TENT TENT TENT TENT TENT	: 1. FORM RIPT A GG C AGG C AGG TT: C ACC Thi	.1522 ATION: ION: G CTG C CAC F GIF 40 ACT F The	SEQ : TGC : TGC : Cys : GAAA : Glu : 25 : GTC : CTA :	TGT Cys 10 ATC Ile ACA Thr TTT Phe	GTG Val CAT His GTC Val TGT Cys	CTG Leu GCC Ala TAT Tyr GCA Ala 60	CGA Arg TAT Tyr 45 TCA Ser	TTC Phe 30 GGG Gly GAT ASP	Cys 15 AGA Arg GTA Vail GCT Ala	AGA Arg CCT Pro AAA Lys	94
GCA Ala GGC Gly GTG Val	GAT Asp GTC Val GGC GLy TGG Trp 50	AGAT	(B) (CD) (CD) (CD) (CD) (CD) (CD) (CD) (CD	NAME, LOCATOR	TION TION TO THE TENT TO THE TENT TENT TENT TENT TENT TENT TENT	: 1. FORM RIPT A GG B GL C AG C AG C Thi S! ACL S!	.1522 ATION: ION: G CTG C CAC F GIF 40 C ACT F Thr	SEQ : TGC : TGC : Cys : GAAA : Glu : 25 : GTC : CTA :	TGT Cys 10 ATC Ile ACA Thr TTT Phe	GTG Val CAT His GTC Val TGT Cys	CTG Leu GCC Ala TAT Tyr GCA Ala 60	CGA Arg TAT Tyr 45 TCA Ser	TTC Phe 30 GGG Gly GAT ASP	Cys 15 AGA Arg GTA Vail GCT Ala	AGA Arg CCT Pro AAA Lys	94 144 197
GCA Ala GGC Val	GAT ASP GTC Val	AGI AGI AGI AGI AGI AGI AGI AGI AGI AGI	(B) (C) (C) (C) (C) (C) (C) (C) (C) (C) (C	NAME LOCATOR CONTROL OF THE CONTROL OF T	TION TION TO THE TOTAL TO THE TOTAL TO THE TOTAL TO THE TOTAL TOTA	: 1. FORM RIPT A GGG Se G Se Le CACO SE LE C	.1522 ATION: ION: G CTG G TGG G TGG G TGG T GI T Thr T AAT A AAN	SEQ TGC Cys GAAGE GAAGE GTC CTA CTA CTA GTT Val	TGT Cys 10 ATC Ile ACA Thr TTT Phe TGG Trp	GTG Val CAT His GTC Val TGT Cys	CTG Leu GCC Ala TAT Tyr GCA Ala 60 ACA Thr	CGA Arg TAT Tyr 45 TCA Ser CAT His	TTC Phe 30 GGG GLY GAT ASP	Cys 15 AGA Arg GTA Val GCT Ala TGT Cys	GLY AGA Arg CCT Pro AAA Lys GTA Val 80 GAA	94 144 197
GCA Ala GGC Val	GAT ASP GTC Val	AGI AGI AGI AGI AGI AGI AGI AGI AGI AGI	(B) (C) (C) (C) (C) (C) (C) (C) (C) (C) (C	NAME LOCATOR CONTROL OF THE CONTROL OF T	TION TION TO THE TOTAL TO THE TOTAL TO THE TOTAL TO THE TOTAL TOTA	: 1. FORM RIPT A GGG Se G Se Le CACO SE LE C	.1522 ATION: G CTG Y Lec C CACG T GGT 40 C ACT T Thr	SEQ TGC Cys GAAGE GAAGE GTC CTA CTA CTA GTT Val	TGT Cys 10 ATC Ile ACA Thr TTT Phe TGG Trp	GTG Val CAT His GTC Val TGT Cys	CTG Leu GCC Ala TAT Tyr GCA Ala 60 ACA Thr	CGA Arg TAT Tyr 45 TCA Ser CAT His	TTC Phe 30 GGG GLY GAT ASP	Cys 15 AGA Arg GTA Val GCT Ala TGT Cys	GLY AGA Arg CCT Pro AAA Lys GTA Val 80 GAA	96 144 197 244
GCA Ala GGC Gly Val GCA Ala 65 CCC Pro	GAT ASP GTC Val GGC Gly TGG Trp 50 TAT Tyr ACA Thr	AGAI AASP	(B) (C) (C) (C) (C) (C) (C) (C) (C) (C) (C	NAME, LOCA:	TION TION TO THE TOTAL TO THE T	: 1. FORM RIPT A GG G G G C AGG C AGG This CAA: GG! AA1	.1522 ATION: ION: G CTG G TGG G TGG G TGG T GI T Thr T AAT A AAN	SEQ TGC Cys GAAA GGC CYS GGC CYS GGC CTA Leu GTT Val	TGT Cys 10 ATC Ile ACA Thr TTT Phe TGG Trp	CAT His GTC Val TGT Cys GCC Ala 75	CTG Leu GCC Ala TAT Tyr GCA Ala 60 ACA Thr	CGA Arg TAT Tyr 45 TCA Ser CAT His AAT ASD	TTC Phe 30 GGG Gly GAT Asp GCC Ala GTA Val	Cys 15 AGA Arg GTA Val GCT Ala TGT Cys ACA Thr GGG	GLY AGA Arg CCT Pro AAA Lys GTA Val 80 GAA GLU GAT	96 144 197 244

			r Le			T CAA p Glr		r Leu					Lys			384
		u Cy				A AAT u Asn 135	Cys					Ala				432
	r As					A ACG y Thr D										480
					r Thi	A AGC - Ser				Glu						528
				Ly:		GAT ASP			Pro							576
			g Lei			TGT Cys		Thr								624
		s Ile				CCA Pro 215										672
	Phe					TGT Cys										720
					Ser	ACA Thr										768
				Gln		CTG Leu										816
			Arg			AAT Asn										864
		Leu				GTA Vai 295										912
Asn 305	Thr	Arg	Lys	Ser	11e 310	CAT His	Ile	Gly	Pro	Gly 315	Arg	Ala	Phe	Tyr	Thr 320	960
ACA Thr	GGA Gly	GAA Glu	ATA Ile	ATA Ile 325	GGA Gly	GAT Asp	ATA	AGA	CAA Gln 330	GCA Ala	CAT His	TGT Cys	AAC Asn	ATT Ile 335	AGT Ser	1008
						ACT Thr	Leu									1056
						ACA /										1104
Asp					Met	CAC / His ! 375				Cys						1152
TAC	TGT	AAT	TCA	ACA	CAA	CTG 1	ITT .	AAT	AGT	ACT	TGG	AAT	AAT	AAT	ACT	1200

(2)	INFO	RMAT	ION	FOR	SEQ	ID N	0:16	:								
			AGA Arg 500				Gln				T GA	GCGG	ccga	:		1532
TAT Tyr	AAA Lys	TAT Tyr	AAA Lys	GTA Val 485	GTA Val	AFA Lys	ATT	Glu	CCA Pro 490	TTA Leu	GGA Gly	GTA Val	GCA Ala	CCC Pro 495	ACC	:488
TTC Phe 465	AGA Arg	CCT Pro	GGA Gly	GGA Gly	GGA Gly 470	GAT ASP	ATG Met	AGG Arg	GAC ASP	AAT Asn 475	TGG Trp	AGA Arg	AGT Ser	GAA GLU	TTA Leu 480	1440
CTG Leu	CTA Leu 450	Leu	ACA Thr	AGA Arg	GAT ASP	GGT Gly 455	GGT Gly	ATT	AAT ASD	GAG Glu	AAT Asn 460	GGG Gly	ACC Thr	GAG Glu	ATC Ile	1392
SCC Ala	CCT Pro	Pro 435	ATC	AGA Arg	GGA Gly	CAA Gln	ATT Ile 440	AGA Arg	TGT Cys	TCA Ser	TCA Ser	AAT Asn 445	ATT Ile	ACA Thr	GGG Gly	1344
ATA	LYS	CAA Glr	ATT 11e 420	lle	AAC Asn	ATG Met	TG6 Trp	GLn 425	GAA Glu	GTA Val	GGA Gly	AAA Lys	GCA Ala 430	ATS Het	TAT Tyr	:2 <del>9</del> 6
GAA	GGC	Ser	AAT Asn	AAC ASF 405	Thr	GAA Glu	GGA Gly	AAT Asn	ACT Thr 410	lle	ACA Thr	CTC L <b>eu</b>	CCA Pro	TGC Cys 415	AGA Arg	1248
1 yr 385		s Ası	n Ser	Thi	Glr 390	i Leu	Phe	. Asn	Ser	Thr 395		ASN	Asn	Asn	Thr 400	

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 507 amino acids
    (B) TYPE: amino acid

  - (D) TOPOLOGY: Linear
- (ii) MOLECULE TYPE: protein
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

Ala Val Phe Val Ser Pro Ser Gln Glu Ile His Ala Arg Phe Arg Arg 20 25 30

Gly Gly Arg Val Glu Lys Leu Trp Val Thr Val Tyr Tyr Gly Val Pro 35 40 45

Val Trp Lys Glu Ala Thr Thr Thr Leu Phe Cys Ala Ser Asp Ala Lys
50 55 60

Ala Tyr Asp Thr Glu Val His Asn Val Trp Ala Thr His Ala Cys Val 65 70 75 80

Pro Thr Asp Pro Asn Pro Gln Glu Val Val Leu Glu Asn Val Thr Glu 85 90 95

His Phe Asn Het Trp Lys Asn Asn Het Val Glu Gln Het Gln Glu Asp 100 105 110

Ile Ile Ser Leu Trp Asp Gln Ser Leu Lys Pro Cys Val Lys Leu Thr 115 120 125

Pro Leu Cys Val Thr Leu Asn Cys Lys Asp Val Asn Ala Thr Asn Thr 130 135 140

Thr Asn Asp Ser Glu Gly Thr Met Glu Arg Gly Glu Ile Lys Asn Cys 145 150 155 160 Ser Phe Asn Ile Thr Thr Ser Ile Arg Asp Glu Val Gln Lys Glu Tyr 165 170 175 Ala Leu Phe Tyr Lys Leu Asp Val Val Pro Ile Asp Asn Asn Asn Thr 180 185 190 Ser Tyr Arg Leu Ile Ser Cys Asp Thr Ser Val Ile Thr Gln Ala Cys 195 200 205 Pro Lys Ile Ser Phe Glu Pro Ile Pro Ile His Tyr Cys Ala Pro Ala 210 215 220 Gly Phe Ala Ile Leu Lys Cys Asn Asp Lys Thr Phe Asn Gly Lys Gly 225 230 235 240 Pro Cys Lys Asn Val Ser Thr Val Gln Cys Thr His Gly Ile Arg Pro 245 250 255 Val Val Ser Thr Gln Leu Leu Leu Asn Gly Ser Leu Ala Glu Glu Glu Glu 260 ... 265 270 Val Val lie Arg Ser Asp Asn Phe Thr Asn Asn Ala Lys Thr lie lie 275 280 285 , Val Gin Leu Lys Glu Ser Val Glu Ile Asn Cys Thr Arg Pro Asn Asn 290 295 300 Asn Thr Arg Lys Ser Ile His Ile Gly Pro Gly Arg Ala Phe Tyr Thr 305 310 315 320 Thr Gly Glu Ile Ile Gly Asp Ile Arg Gln Ala His Cys Asn Ile Ser 325 330 335 Arg Ala Lys Trp Asn Asp Thr Leu Lys Gln Ile Val Ile Lys Leu Arg 340 345 350 Glu Gln Phe Glu Asn Lys Thr Ile Val Phe Asn His Ser Ser Gly Gly 355 360 365 Asp Pro Glu Ile Val Met His Ser Phe Asn Cys Gly Gly Glu Phe Phe 370 375 380 Tyr Cys Asn Ser Thr Gln Leu Phe Asn Ser Thr Trp Asn Asn Asn Thr 385 390 395 400 Glu Gly Ser Asn Asn Thr Glu Gly Asn Thr Ile Thr Leu Pro Cys Arg 405 410 415 Ile Lys Gin Ile Ile Asn Met Trp Gin Glu Val Gly Lys Ala Met Tyr 420 425 430 Ala Pro Pro Ile Arg Gly Gln Ile Arg Cys Ser Ser Asn Ile Thr Gly
435 440 445 Leu Leu teu Thr Arg Asp Gly Gly Ile Asn Glu Asn Gly Thr Glu Ile 450 455 460 Phe Arg Pro Gly Gly Gly Asp Met Arg Asp Asn Trp Arg Ser Glu Leu 465 470 475 480 Tyr Lys Tyr Lys Val Val Lys Ile Glu Pro Leu Gly Val Ala Pro Thr 485 490 495 Lys Ata Lys Arg Arg Val Val Gin Arg Giu Lys 500 505

•	71	INFORMATIO	N FOR	SED	10	NO · 1	7.
٠,		INFUNDATION	- · ·	364	10	NU:	

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 1484 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

### (11) MOLECULE TYPE: DNA (genomic)

### (ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 1..1474
- (D) OTHER INFORMATION:

### (xi) SEQUENCE DESCRIPTION: SEO ID NO: 17:

(	xí)	SEGUE	NCE	DESC	RIPT	ON:	SEO	ID A	10: 17	<b>7:</b>					
ATG G Met A 1	AT GI	CA A1 la Me	t Ly	G AGA S Arg 5	GGC GLy	CTC Leu	TGC Cys	TGT Cys	. Vai	CTC Lei	CTG J Leu	CTG Leu	TGT Cys 15	Gly	48
GCA G	TC T1 ai Pi	ne Va	T TCI	G CCC	AGE Ser	CAG Gin	GAA Glu 25	Ite	CAT	GCC	CGA Arg	Phe 30	VLA	AGA	96
GGC GI Gly A	la Ar 3	g Th	r Gli	ı Lys	Leu	1rp 40	Val	Thr	Val	Tyr	1yr 45	Gly	Val	Pro	144
	rp Ly 50	s Gl	J Ala	i inr	Thr 55	Thr	Leu	Phe	Cys	Ala 60	Ser	Asp	Ala	Lys	192
GCA TA Ala:Ty 65	r As	p Thi	- Głu	70	His	Asn	Val	Тгр	Ala 75	Thr	His	Ala	Cys	Val 80	240
CCC AC Pro Th	r As	p Pro	ASN 85	Pro	Gln	Glu	Val	90	Leu	Vat	Asn	Val	Thr 95	Glu	288
AAT TT ASD Ph	e Asi	100	1rp	Lys	ASN	ASP	Met 105	Val	Glu	Gln	Met	His 110	Glu	Asp	336
ATA ATI	e Ser 115	Leu	Trp	Asp	Gln	Ser 120	Leu	Lys	Pro	Cys	Val 125	Lys	Leu	Thr	384
Pro Les 130	u Cys )	Val	Ser	Leu	135	Cys	Thr	Asp	Leu	Gly 140	Asn	Ala	Thr	Asn	432
ACC AA1 Thr Asn 145	) Ser	Ser	Asn	Thr / 150	Asn :	Ser :	Ser	Ser	Gly 155	Glu	Het	Het	Met	Glu 160	480
AAA GGA Lys Gly	Glu	Ile	165	Asn (	Cys S	ier í	Phe /	170	ile :	Ser	Thr	Ser	11e 175	Arg	528
GGT AAG Gly Lys	Val	Gln 180	Lys (	ilu 1	YF A	la f	Phe 1 185	he 1	[yr	Lys	Leu	Asp 190	lie	It€	576
CCA ATA Pro Ile	GAT ASP 195	AAT Asn	GAT A Asp 1	CT A	hr S	GC 1 er 1 00	YAT A	CG 1	ITG /	Ihr	AGT Ser 205	TGT Cys	AAC Asn	ACC Thr	624

TCA Ser	GTC Val 210	Ιle	ACA Thr	CAG Gln	GCC	TGT Cys 215	CCA Pro	AAG Lys	GTA Val	TCC Ser	111 Phe 220	GAG Glu	CCA Pro	ATT !le	CCC Pro	672
ATA 11e 225	CAT His	TAT Tyr	TGT Cys	GCC	CCG Pro 230	GCT Ala	GGT Gly	TTT Phe	GCG	ATT Ile 235	CTA Leu	AAA Lys	TGT Cys	AAT Asn	AAT Asn 240	<b>720</b>
AAG Lys	ACG Thr	TTC Phe	AAT Asn	GGA Gly 245	Thr	GGA Gly	CCA Pro	TGT Cys	ACA Thr 250	AAT Asn	GTC Val	AGC Ser	ACA Thr	GTA Val 255	CAA Gln	768
TGT Cys	ACA Thr	CAT His	GGA Gly 260	ATT	AGG Arg	CCA Pro	GTA Val	GTA Val 265	TCA Ser	ACT Thr	CAA Gln	CTG Leu	CTG Leu 270	TTG Leu	AAT Asn	816
GGC Gly	AGT Ser	CTA Leu 275	GCA Ala	GAA Glu	GAA Glu	GAG Glu	GTA Vai 280	GTA Val	ATT	AGA <sup>.</sup> Arg	TCT Ser	GCC Ala 285	AAT Asn	TTC Phe	ACA Thr	864
GAC Asp	AAT Asn 290	Ala	AAA Lys	ACC Thr	ATA	ATA Ile 295	GTA Val	CAG Gin	CTG Leu	AAC Asn	CAA Gln 300	TCT Ser	GTA Val	GAA Glu	ATT Ile	912
AAT Asn 305	TGT Cys	ACA Thr	GGT Gly	GCT	GGA Gly 310	CAT His	TGT Cys	AAC Asn	ATT Ile	AGT Ser 315	AGA Arg	GCA Ala	AAA Lys	TGG Trp	AAT Asn 320	<b>960</b>
GCC	ACT Thr	TTA Leu	AAA Lys	CAG Gln 325	ATA Ile	GET Ala	AGC Ser	AAA Lys	TTA Leu 330	AGA Arg	GAA Glu	CAA Gln	TTT Phe	GGA Gly 335	AAT Asn	1008
AAT Asn	AAA Lys	ACA Thr	ATA Ile 340	ATC Ile	TTT Phe	AAG Lys	CAA Gln	TCC Ser 345	TCA Ser	GGA Gly	GGG Gly	GAC Asp	CCA Pro 350	GAA Glu	ATT Ile	<b>1056</b>
				TTT Phe												1104
				AAT Asn												1152
GGG Gly 385	TCA Ser	AAT Asn	AAC Asn	ACT Thr	GAA Glu 390	GGA Gly	AGT Ser	GAC Asp	ACA Thr	ATC I te 395	ACA Thr	CTC Leu	CCA Pro	TGC Cys	AGA Arg 400	1200
				ATA Ile 405											Tyr	1248
GCC Ala	CCT Pro	CCC Pro	ATC I le 420	AGC Ser	GGA	CAA Gln		AGA Arg 425					~	The	GCG	1296
CTG Leu	CTA Leu	TTA Leu 435	ACA Thr	AGA Arg	GAT Asp	GGT Gly	GGT Gly 440	AAT Asn	AAC Asn	AAC Asn	AAT Asn	GGG Gly 445	Ser	GAG	ATC Ile	1344
TTC Phe	AGA Arg 450	CCT Pro	GGA Gly	GGA Gly	Gly	GAT ASP 455	ATG Met	AGG Arg	GAC Asp	AAT Asn	TGG Trp 460	Arg	AGT Ser	GN	L TTA	1392
TAT Tyr 465	AAA Lys	TAT Tyr	AAA Lys	Val	GTA Val 470	AAA Lys	ATT Ile	GAA Glu	CCA Pro	TTA Leu 475	GGA Gly	GTA Val	GCA Ala	Pre	ACC Thr 480	1440
AAG	GCA	AAG	AGA	AGA	GTG	GTG	CAS	AGA	GAA	**	T G	AGCG	GCCG	iC .		1484

Lys Ala Lys Arg Arg Val Val Gin Arg Giu Lys 485 490

- (2) INFORMATION FOR SEQ ID NO:18:
  - (i) SEQUENCE CHARACTERISTICS:
    - (A) LENGTH: 491 amino acids
    - (8) TYPE: amino acid
      (D) TOPOLOGY: linear
  - (ii) MOLECULE TYPE: protein
  - (xi) SEQUENCE DESCRIPTION: SEO ID NO:18:

Met Asp Ala Met Lys Arg Gly Leu Cys Cys Val Leu Leu Leu Cys Gly 1 5 10 15

Ala Val Phe Val Ser Pro Ser Gln Glu Ile His Ala Arg Phe Arg Arg 20 25 30

Gly Ala Arg Thr Glu Lys Leu Trp Val Thr Val Tyr Tyr Gly Val Pro 35 40 45

Val Trp Lys Glu Ala Thr Thr Thr Leu Phe Cys Ala Ser Asp Ala Lys
50 60

Ata Tyr Asp Thr Glu Val His Asn Val Trp Ata Thr His Ata Cys Val 65 70 75 80

Pro Thr Asp Pro Asn Pro Gln Glu Val Val Leu Val Asn Val Thr Glu 85 90 95

Asn Phe Asn Het Trp Lys Asn Asp Met Val Glu Gln Met His Glu Asp 100 105 110

Ite Ite Ser Leu Trp Asp Gln Ser Leu Lys Pro Cys Val Lys Leu Thr 115 120 . 125

Pro Leu Cys Val Ser Leu Lys Cys Thr Asp Leu Gly Asn Ala Thr Asn 130 135 140

Thr Asn Ser Ser Asn Thr Asn Ser Ser Ser Gly Glu Met Met Met Glu 145 150 150 160

Lys Gly Glu ile Lys Asn Cys Ser Phe Asn ile Ser Thr Ser ile Arg 165 170 175

Gly Lys Val Gln Lys Glu Tyr Ala Phe Phe Tyr Lys Leu Asp Ile Ile 180 185 190

Pro Ile Asp Asn Asp Thr Thr Ser Tyr Thr Leu Thr Ser Cys Asn Thr 195 200 205

Ser Val Ile Thr Gln Ala Cys Pro Lys Val Ser Phe Glu Pro Ile Pro 210 215 220

Ile His Tyr Cys Ala Pro Ala Gly Phe Ala Ile Leu Lys Cys Asn Asn 225 230 235 240

Lys Thr Phe Asn Gly Thr Gly Pro Cys Thr Asn Val Ser Thr Val Gln 245 250 255

Cys Thr His Gly Ile Arg Pro Val Val Ser Thr Gln Leu Leu Asn 260 265 270

Gly Ser Leu Ala Glu Glu Glu Val Val Ile Arg Ser Ala Asn Phe Thr 275 280 285

As	p As 29		aL	ys 1	ihr	Ile	1 l e 295		Glr	1 Lei	J Asn	300		Val	Glu	He	
As:		s Th	ır G	ly #		Gly 310		Cys	s Asr	ı Ite	Ser 315		Ala	Lys	Trp	Asn 320	
AL	a Th	r Le	u L		iln 25	Ile	Ala	Ser	Lys	330	Arg	Glu	Gln	Phe	Gly 335		
Asi	n Ly	s Th		le 1 40	i e	Phe	Lys	Glr	Ser 345		Gly	Gly	Asp	Pro 350		Ile	
Val	Th	r Hi 35		er P	he i	Asn	Cys	Gl y 360		Glu	Phe	Phe	Tyr 365	Cys	Asn	Ser	
Thr	Gl1 370		u Pl	ne A	sn :	Ser	Thr 375	Trp	Phe	Asn	Ser	Thr 380	•	Ser	Thr	Glu	
Gly 385		- As	n As	n T	hr (	3lu 590	Gly	Ser	Asp	Thr	11e 395	Thr	Leu	Pro	Cys	Arg 400	
ile	Ly:	; Gli	n Ph		le / 05	lsn	Het	îrp	Gln	Glu 410	Val	Gly	Lys	Ala	Met 415	Tyr	
Ala	Pro	Pr	0 I I	.e Se 20	er C	ily	Gln	lle	Arg 425	Cys	Ser	Ser	Asn	11e 430	Thr	Gly	
		43!	5					440			Asn		445			•	
	450	1					455	-			Asn	460					
465					4	70					Leu 475	Gly	Val	Ala	Pro	Thr 480	
				48	5				Arg	6lu 490	Lys						
	(ii.	) SE (( ( ( ( ( ( ( ( ( ( ( ( ( ( ( ( ( (	QUEI A) I B) I C) ! D) I LECL ATUR A) R B) L	NCE LENG TYPE STRAI TOPO JLE IAME, OCA I	CHAI TH: : NDEI LOGY TYPE /KCEY	RACT 144 DNES 7: (	TERI 48 beic 55: Line DNA	STIC ase acid sing ar (gen	S: pair l le conic	:)	: 19:						
ATG Met	GAT	GCA	ATG	MG	AG Ar	A G	GG (	TC '	TGC	TGT	STG	CTG					48
GCA Ala				Ser						ATC					AGA Arg		96
GGC Gly							eu T										144

							467	CT.	***	767		T.C.4	CA7	CCT		192
Val	7 rp 7 rp 50	LYS	GL	Ala	ACC Thr	Thr 55	Thr	Leu	Phe	Cys	Ala 60	Ser	Asp	Ala	Lys ·	
GCA Ala 65	Tyr	GAT Asp	ACA Thi	GAG	GTA Val 70	His	AAT Asn	GTT Val	TGG Trp	GCC Ala 75	ACA Thr	CAT	GCC Ala	TGT Cys	GTA Val 80	240
					CCA Pro											288
CAT His	TTT Phe	AAC ASN	ATG Met 100	Trp	AAA Lys	AAT Asn	AAC Asn	ATG Met 105	Val	GAA Glu	CAG Gln	ATG Met	CAG Gln 110	GAG Glu	GAT Asp	336
ATA !le	ATC 1le	AGT Ser 115	Leu	TGG Trp	GAT Asp	CAA Gln	AGC Ser 120	Leu	AAG Lys	CCA Pro	TGT Cys	GTA Val 125	AAA Lys	TTA Leu	ACC Thr	384
		Cys			TTA Leu											432
	Asn				GGA Gly 150											480
					ACA Thr											528
				Lys	CTT Leu											576
					AGT Ser											624
					GAG Glu											672
					AAG Lys 230										GGA Gly 240	720
															CCA Pro	768
GTA Val	GTA Val	TCA Ser	ACT Thr 260	CAA Gln	CTG L <b>e</b> u								•	GAA Glu	GAG Glu	816
GTA Val	GTA Val	ATT 11e 275	AGA Arg	TCT Ser	GAC Asp	Asn	TTC Phe 280	ACG Thr	AAC Asn	AAT Asn	GCT Ala	AAA Lys 285	ACC Thr	ATA I le	ATA 1le	864
Val	CAG Gln 290	CTG L <b>eu</b>	AAA Lys	GAA Glu	Ser	GTA Val 295	GAA Glu	ATT Ile	AAT Asn	TGT Cys	ACA Thr 300	GGT Gly	GCT Ala	GGA Gly	CAT His	912
TGT Cys 305	AAC Asn	ATT Ile	AGT Ser	Arg	GCA Ala 310	AAA Lys	TGG Trp	AAT Asn	GAC Asp	ACT Thr 315	TTA Leu	AAA Lys	CAG Gln	ATA	GTT Val 320	960
ATA	***	ATT	AGA	GAA	CAA	TTT	GAG	AAT	***	ACA	ATA	GTC	777	AAT	CAC	1008

110	e Ly	s Le	u Ar	g Gli 329		n Phe	GLL	J AST	330		Ile	Val	Phe	Asn 335	His	
				A YE		A GAA D Glu			Met					Cys		1056
			e Pho			AAT Asn		Thr								1104
		1 As				TCA Ser 375						Asn				1152
						CAA Gln										1200
						CCC Pro										1248
				Leu		TTA Leu										1296
GGG Gly	ACC Thr	GAG Glu 435	ATC Ile	TTC Phe	AGA Arg	CCT Pro	GGA Gly 440.	Gly	GGA Gly	GAT Asp	ATG Met	AGG Arg 445	GAC ASP	AAT ASD	TGG Trp	1344
AGA Arg	AGT Ser 450	GAA Glu	TTA Leu	TAT Tyr	Lys	TAT Tyr 455	AAA Lys	GTA Val	GTA Val	AAA Lys	ATT Ile 460	GAA Glu	CCA Pro	TTA Leu	GGA Gly	1392
GTA Val 465	GCA Alb	CCC Pro	ACC Thr	Lys	GCA Ala 470	AAG Lys	AGA Arg	AGA Arg	Val	GTG Val 475	CAA Gln	AGA Arg	GAA Glu	AAA Lys	TG	1439
AGCG	GCCC	iC														1448

## (2) INFORMATION FOR SEQ ID NO:20:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 479 amino acids
  - (B) TYPE: amino acid
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: protein
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

Met Asp Ala Met Lys Arg Gly Leu Cys Cys Val Leu Leu Leu Cys Gly
1 5 10 15

Ala Val Phe Val Ser Pro Ser Gln Glu Ile His Ala Arg Phe Arg Arg 20 25 30

Gly Gly Arg Val Glu Lys Leu Trp Val Thr Val Tyr Tyr Gly Val Pro 35 40 45

Val Trp Lys Glu Ala Thr Thr Thr Leu Phe Cys Ala Ser Asp Ala Lys 50 55 60

Ala Tyr Asp Thr Glu Val His Asn Val Trp Ala Thr His Ala Cys Val 65 70 75 80

Pro Thr Asp Pro Ash Pro Gln Glu Val Val Leu Glu Ash Val Thr Glu

				85	5				90	)				95	
His	Phe	e Asr	100	t Trp D	lys	Ası	n Asr	Met 105		Glu	Gln	Net	Gln 110	Glu	Asp
Ilе	110	Ser 115		u Trp	) ASE	Glr	1 Ser 120		l Lys	Pro	Cys	Val 125	Lys	Leu	Thr
Pro	130		, va	l Thr	. Len	135		Lys	Asp	Val	Asn 140	Ala	Thr	Asn	Thr
1hr 145		ASF	Ser	- Głu	150		Met	Glu	Arg	Gly 155	Glu	Ile	Lys	Asn	Cys 160
Ser	Phe	Asr	i i le	165		Ser	Ile	Arg	170		Val	Gln	Lys	Glu 175	Туг
Ala	Leu	Phe	180	Lys	leu	Asp	Val	Va l 185	Pro	ile	Asp	Asn	190	Asn	Thr
Ser	Tyr	Arg 195		ı Ile	Ser	Cys	200		Ser	Val	lle	1hr 205	Gln	Ala	Cys
	210	+		Phe		215					220				
225				Leu	230					235					240
		•		245					250			_		255	
			260					265					270		
		275		Ser			280					285			
	290			Glu		295				·	300	·		٠	
305				Arg	310		•		·	315					320
	·			325					330					335	
		•	340	Asp				345					350	-•-	
•		355		Tyr	-		360					365			•
	370			Glu		375	Asn				380				
385					390					395					400
•				Ala 405					410					415	
			420	Leu				425					430		
•		435		Phe -			440					445			
1	C	GI.	1 41	Tvr :	I VS '	TVP	IVE	Val	Val	i ve	110	Glas	Pro	1	FIV

	450	ı				455					460						
va I 465		Pro	Thi	r Lys	470		Arg	Arg	Val	Val 475	űln	Arg	Glu	Lys			
(2)	INF	ORMA	TIO	N FOR	SEC	DI G	NO:2	1:									
	(i	(	A) 1 B) 1 C) 5	NCE ( LENG1 TYPE: STRAN	iH: 1 nuc IDEDN	484 leic ESS:	base aci sin	pa i d	FS								
	(ii	) MO	LECL	JLE T	YPE:	DNA	( ge	nomi	c)								
	(ix	(	B) L	RE: IAME/ OCAT OTHER	1 ON :	1		:									
	(xi	) SE	QUEN	ICE D	ESCR	IPTI	ON:	SEQ	ID NO	0:21	:				•		
ATG Met 1	Asp	GCA	ATG Met	AAG Lys 5	Arg	GGG	CTC Leu	TGC Cys	TGT Cys 10	GTG Val	CTG Leu	C16 Leu	CTG L <del>e</del> u	TGT Cys 15	GGA Gly	ž	8
				TCG Ser												(	96
GC ily	GCC Ala	AGA Arg 35	Thr	GAA Glu	AAA Lys	TTG Leu	TGG Trp 40	GTC Val	ACA Thr	GTC Val	TAT Tyr	TAT Tyr 45	GGG Gly	GTA Val	CCT Pro	14	44
		Lys		GCA Ala												. 1	92
iCA la 65	TAT Tyr	GAT ASP	ACA Thr	GA6 Glu	GTA Val 70	CAT	AAT Asn	GTT Val	TGG Trp	GEC Ala 75	ACA Thr	CAT His	GCC Ala	TGT Cys	GTA Val 80	2	40
CC	ACA Thr	GAC Asp	CCC Pro	AAC Asn 85	CCA Pro	CAA Gln	GAA Glu	GTA Val	GTA Val 90	TTG Leu	GTA Val	AAT ASD	GTG Val	ACA Thr 95	GAA Glu	2	88
AT sn	111 Phe	AAC Asn	ATG Met 100		AAA Lys	AAT Asn	GAC Asp	ATG Met 105	GTA Val	GAA Glu	CAG Gln	ATG Met	CAT His 110	GAG Glu	GAT Asp	3	36
TA le	ATC Ile	AGT Ser 115	ITA Leu	TGG Trp	GAT Asp	CAA Gln	AGC Ser 120	CTA	AAG Lys	CCA Pro	TGT Cys	GTA Val 125	AAA Lys	TTA Leu	ACC Thr	3	84
															AAT Asn	4	32
CC hr 45	AAT Asn	AGT Ser	AGT Ser	AAT Asn	ACC Thr 150	AAT Asn	AGT Ser	AGT Ser	AGC Ser	GGG Gly 155	GAA Glu	ATG Met	ATG Met	ATG Het	GAG Glu 160	4	480
AA ys	GGA Gly	GAG Glu	ATA Ile	AAA Lys 165	AAC Asn	TGC Cys	TCT Ser	TTC Phe	AAT Asn 170	ATC Ile	AGC Ser	ACA Thr	AGC Ser	ATA Ile 175	AGA	• • •	528
GT	AAG	GTG	CAG	AAA	GAA	TAT	GCA	TTT	TTT	TAT	AAA	CTT	GAT	ATA	ATA	!	576

	180	405		
	180	185	190	
Pro Ile Asp 195		Ser Tyr Thr Leu 1 200	Thr Ser Cys Asn 205	Thr
TCA GTC ATT A Ser Val Ile 1 210	ACA CAG GCC TGT C Thr Gln Ala Cys P 215	ro Lys Val Ser P	TT GAG CCA ATT The Glu Pro Ile 20	CCC 672 Pro ·
ATA CAT TAT T Ite His Tyr C 225	GT GCC CCG GCT G Cys Ala Pro Ala G 230	GT TTT GCG ATT C ly Phe Ala Ile L 235	TA AAA TGT AAT eu Lys Cys Asn	AAT 720 Asn 240
AAG ACG TTC A Lys Thr Phe A	AT GGA ACA GGA C isn Gly Thr Gly P 245	CA TGT ACA AAT G ro Cys Thr Asn V 250	TC AGC ACA GTA al Ser Thr Val 255	CAA 768 Gln
Cys Thr His G	GA ATT AGG CCA G ly Ile Arg Pro Va 60	TA GTA TCA ACT C al Val Ser Thr G 265	AA CTG CTG TTG In Leu Leu Leu 270	AAT. 816 Asn
GGC ACT CTA S Gly Ser Leu A 275	CA CAA GAA GAG G1 la Glu Glu Glu Va 28	si Val ile Arg S	CT GCC AAT TTC er Ala Asn Phe 285	ACA 854 Thr
GAC AAT GCT AA ASP ASP Ala Ly 290	MA ACC ATA ATA GT ys Thr Ile Ile Va 295	ol Gin Leu Asn G	MA TCT GTA GAA In Ser Val Glu DO	ATT 912 Ile
AAT TGT ACA GO Asn Cys Thr GO 305	GT GCT GGA CAT TG Ly Ale Gly His Cy 310	it AAC ATT AGT AG 's Asn Ile Ser Ag 315	rg Ala Lys Trp	AAT 960 Asn 320
GCC ACT TTA AA Ala Thr Leu Ly	A CAG ATA GCT AG 'S Gin lie Ala Se 325	C AAA TTA AGA GA F Lys Leu Arg GI 330	MA CAA TTT GGA . iu Gln Phe Gly . 335	AAT 1008 Asn
AAT AAA ACA AT Asn Lys Thr 11 34	A ATC TIT AAG CA e lie Phe Lys Gi O	A TCC TCA GGA GG n Ser Ser Gly Gi 345	G GAC CCA GAA Y ASP Pro Glu 350	ATT 1056 lie
GTA ACG CAC AG Val Thr His Se 355	T TTT AAT TGT GG r Phe Asn Cys Gly 360	y Gly Glu Phe Ph	C TAC TGT AAT : e Tyr Cys Asn : 365	TCA 1104 Ser
370	T AAT AGT ACT TGG ASN Ser Thr Tr; 375	Phe Asn Ser Th 38	r Trp Ser Thr ( O	Glu
Gly Ser Ash Asr 385	ACT GAA GGA AG1 Thr Giu Gly Ser 390	r Asp Thr Ile Th 395	r Leu Pro Cys	Arg 400
ATA AAA CAA TTT Ile Lys Gin Phe	ATA AAC ATG GTG Ile Asn Met Val 405	CAG GAA GTA GG Gin Glu Val Gi 410	A AAA GCA ATG 1 y Lys Ala Met 1 415	rat 1248 Typ
GCC CCT CCC ATC Ala Pro Pro Ile 420	AGC GGA CAA ATT Ser Gly Gln Ile	AGA TGT TCA TCA Arg Cys Ser Ser 425	A AAT ATT ACA ( r Asn lie Thr ( 430	5G6 1296 Sly
CTG CTA TTA ACA Leu Leu Thr 435	AGA GAT GGT GGT Arg Asp Gly Gly 440	ASD ASD ASD ASI	r GGG Ter GAG A n Gly Ser Glu 1 445	ATC 1344 I le
TTC AGA CCT GGA Phe Arg Pro Gly 450	GGA GGA GAT ATG Gly Gly Asp Met 455	AGG GAC AAT TGE Arg Asp Asn Tre 460	Arg Ser Glu L	ITA 1392 .eu

TAT AAA TAT AAA GTA GTA AAA ATT GAA CCA TTA GGA GTA GCA CCC ACC
Tyr Lys Tyr Lys Val Lys Ile Glu Pro Leu Gly Val Ala Pro Thr
465 470 475 480

AAG GCA AAG AGA AG AGTGGTGCAG AGAGAAAAAT GAGCGGCCGC Lys Ala Lys Arg 1484

#### (2) INFORMATION FOR SEQ ID NO:22:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 484 amino acids
  - (B) TYPE: amino acid
  - (D) TOPOLOGY: Linear
- (ii) MOLECULE TYPE: protein
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

Met Asp Ala Met Lys Arg Gly Leu Cys Cys Val Leu Leu Leu Cys Gly
1 5 10 15

Ata Vat Phe Val Ser Pro Ser Gln Glu Ite His Ata Arg Phe Arg Arg 20 25 30

Gly Ala Arg Thr Glu Lys Leu Trp Val Thr Val Tyr Tyr Gly Val Pro 35 40 45

Val Trp Lys Glu Ala Thr Thr Thr Leu Phe Cys Ala Ser Asp Ala Lys 50 55 60

Ala Tyr Asp Thr Glu Val His Asn Val Trp Ala Thr His Ala Cys Val 65 70 75 80

. Pro Thr Asp Pro Asn Pro Gin Glu Val Val Leu Val Asn Val Thr Glu 85 90 95

Asn Phe Asn Met Trp Lys Asn Asp Met Val Glu Gln Met His Glu Asp 100 105 110

Ile Ile Ser Leu Trp Asp Glm Ser Leu Lys Pro Cys Val Lys Leu Thr 115 120 125

Pro Leu Cys Val Ser Leu Lys Cys Thr Asp Leu Gly Asn Ala Thr Asn 130 135 140

Thr Asn Ser Ser Asn Thr Asn Ser Ser Ser Gly Glu Met Met Glu 145 150 155 160

Lys Gly Glu Ile Lys Asn Cys Ser Phe Asn Ile Ser Thr Ser Ile Arg 165 170 175

Gly Lys Val Gln Lys Glu Tyr Ala Phe Phe Tyr Lys Leu Asp Ile Ile 180 185 190

Pro Ile Asp Asn Asp Thr Thr Ser Tyr Thr Leu Thr Ser Cys Asn Thr 195 200 205

Ser Val Ile Thr Gln Ala Cya Pro Lys Val Ser Phe Glu Pro Ile Pro 210 215 220

lle His Tyr Cys Ala Pro Ala Ely Phe Ala Ile Leu Lys Cys Asn Asn 225 230 235 240

Lys Thr Phe Asn Gly Thr Gly Pro Cys Thr Asn Val Ser Thr Val Gln 245 250 255

Cys Thr His Gly Ile Arg Pro Val Val Ser Thr Gln Leu Leu Leu Asn

			260	)				265					270			
Gly	Ser	275		Glu	Gli	Glu	val 280		lle	Arg	Ser	Ala 285	Asn	Phe	Thr	
Asp	290		Lys	Thr	He	1 le 295		Gln	Leu	ASD	Gln 300		Val	Glu	Ite	
Asn 305		Thr	Gly	Ala	Gly 310		Cys	Asn	Ιle	Ser 315	Arg	Ala	Lys	îrp	Asn 320	
Ala	Thr	Leu	Lys	Gln 325	Ile	Ala	Ser	Lys	L eu 330	Arg	Glu	Gin	Phe	Gly 335	Asn	
Asn	Lys	Thr	I l e 340	Ile	Phe	Lys	Gln	Ser 345	Ser	Gly	Gly	Asp	Pro 350	Glu	Ile	
val	Thr	His 355	Ser	Phe	Asn	Cys	Gly 360		Glu	Phe	Phe	Туг 365	Cys	Asn	Ser	
Thr	Gln 370	Leu	Phe	Asn	Ser	Thr 375	Trp	Phe	Asn	Ser	Thr 380	Trp	Ser	Thr	Glu	
Gly 3 <b>85</b>	Ser	Asn	Asn	Thr	Glu 390	Gly	Ser	Asp	Thr	ile 395	Thr	Leu	Pro	Cys	Arg 400	
Ile	Lys	Gln	Phe	11e 405	Asn	Met	Val	Gln	Glu 410	Val	Gly	Lys	Ala	Met 415	Tyr	
Ala	Pro	Pro	1 le 420	Ser	Gly	Gln	Ile	Arg 425	Cys	Ser	Ser	Asn	11e 430	Thr	Gly	
Leu	Leu	Leu 435	Thr	Arg	Asp	Gly	Gly 440		Asn	Asn	Asn	Gly 445	Ser	Glu	Ile	
	450			Gly		455					460					
Tyr 465	Lys	Туг	Lys	Val	Val 470	Lys	He	Glu		Leu 475	Gly	Val	Ala	Pro	Thr 480	
Lys	Ala	Lys	Arg													
(2)	INFO	RMAT	ION	FOR	SEQ	ID N	0:23	:								
٠	(i)	(A (B (C	) LE ) TY ) ST	E CHA NGTH: PE: I RANDI POLOI	: 14 nucl EDNE	48 b eic SS:	ase acid sing	pair	5							
	(ii)	MOL	ECULI	E TYI	PE: 1	DNA	(gen	omi c	)							
	(ix)	FEAT	TURE:	:												

#### (xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

(A) NAME/KEY: CDS
(B) LOCATION: 1..1438
(D) OTHER INFORMATION:

ATG GAT GCA ATG AAG AGA GGG CTC TGC TGT GTG CTG CTG CTG TGT GGA

Met Asp Ala Met Lys Arg Gly Leu Cys Cys Val Leu Leu Leu Cys Gly

1 5 15

GCA GTC TTC GTT TCG CCC AGC CAG GAA ATC CAT GCC CGA TTC AGA AGA

Ala Val Phe Val Ser Pro Ser Gln Glu Ile His Ala Arg Phe Arg Arg

20 25 30

		y Ar		TA GA al Gi				val								144
-		P Ly		AA GC			· Thi					Ser				192
	a Ty			A GA		l His					Thr					240
				C AA O As B	n Pro					Leu						288
				G TG T Tr					Val							336
			r Le	A TG				Leu								384
		ı Cy		T ACT			Cys					Ala				432
	Asn			C GAC		Thr										480
				C ACC Thr 165	Thr											528
				Lys												576
AGE Ser	TAT	AGG Arg 195	Leu	ATA Ile	AGT Ser	TGT Cys	GAC Asp 200	ACC Thr	TCA Ser	GTC Val	ATT	ACA Thr 205	EAG Gln	GCC Ala	TGT Cys	624
				††††												672
				CTA Leu												720
CCA Pro	TGT Cys	AAA Lys	AAT	GTC Vel 245	Ser	ACA Thr	GTA Val	CAA Gln	TGT Cys 250	ACA Thr	CAT His	GGA Gly	ATT	AGG Arg 255	Pro	768
				CAA Gln										Glu		816
	Val			TCT Ser		Asn										864
					Ser										CAT His	912
GT	AAC	ATT	AGT	AGA	GCA	AÄA	TGE	AAT	GAC	ACT	TTA	AAA	CAG	ATA	GTT	960

305		n il	e se	T AT	310		ז ון	D ASI	n Asp	315		I LYS	Gln	lie	320	
ATA !!e	A AA	A TT	A AG	A GAZ 9 Gli 32	ı Glr	1 TTI	GAI	G AA1 J Asr	AAA Lys 330	Thr	ATA Ile	GTC	††† Phe	AAT Asn 335	CAC	1008
TCC Ser	Ser	GG	A GG y Gl	G GAO y Asp D	CC#	GAA	AT1	GTA Val 345	Het	CAE His	AGT Ser	TTT Phe	AAT Asn 350	Cys	GGA Gly	1056
GGA Gly	GAA	771 Phe 355	e Pho	TAC Tyr	TGT Cys	AAT Asn	TCA Ser 360	Thr	CAA Gln	CTG Leu	TTT Phe	AAT Asn 365	AGT Ser	ACT Thr	TGG Trp	1104
AAT Asn	AAT Asn 370	Asr	AC1	GAA Glu	GGG	TCA Ser 375	AAT Asn	AAC Asn	ACT Thr	GAA Glu	GGA Gly 380	AAT Asn	ACT Thr	ATC Ile	ACA Thr	1152
CTC Leu 385	CCA Pro	TGC	AGA Arg	ATA   Ile	AAA Lys 390	CAA Gln	ATT	ATA	AAC Asn	ATG Met 395	GTG Val	CAG Gln	GAA Glu	GTA Val	GGA Gly 400	1200
AAA Lys	GCA Ala	ATG Met	TAT	GCC Ala 405	CCT Pro	Pro	ATC	AGA Arg	GGA Gly 410	CAA Gln	ATT Ile	AGA Arg	TGT Cys	TCA Ser 415	TCA Ser	1248
AAT Asn	ATT	ACA Thr	GGG Gly 420	CTG L <b>eu</b>	CTA Leu	TTA Leu	ACA Thr	AGA Arg 425	GAT Asp	GGT Gly	GGT Gly	ATT ile	AAT Asn 430	GAG Glu	AAT Asn	12 <del>96</del>
GGG Gly	ACC Thr	GAG Glu 435	ATC Ile	TTC Phe	AGA Arg	CCT Pro	GGA Gly 440	GGA Gly	GGA Gly	GAT Asp	ATG Met	AGG Arg 445	GAC Asp	AAT Asn	TGG Trp	1344
lrg	AGT Ser 450	GAA Glu	TTA Leu	TAT	Lys	TAT Tyr 455	AAA Lys	GTA Val	GTA Val	Lys	ATT Ile 460	GAA Glu	CCA Pro	TTA Leu	GGA Gly	1392
al 65	GCA Ala	CCC Pro	ACC Thr	AAG Lys	GCA Ala 470	AAG Lys	AGA Arg	AGA Arg	Val	GTG Val 475	CAA Gln	AGA Arg	GAA Glu	AAA Lys	T	1438
AGC	GGCC	GC														1448

### (2) INFORMATION FOR SEQ ID NO:24:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 479 amino acids
    (B) TYPE: amino acid
    (D) TOPOLOGY: linear

#### (ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

Ala Val Phe Val Ser Pro Ser Gln Glu Ile His Ala Arg Phe Arg Arg 20 25 30

Gly Gly Arg Val Glu Lys Leu Trp Val Thr Val Tyr Tyr Gly Val Pro 35 40 45

Val Trp Lys Glu Ala Thr Thr Thr Leu Phe Cys Ala Ser Asp Ala Lys 50 55 60

	l a 65	7,	/r .	Asp	<b>5</b> T	hr	Gl	u V	a ( 70	His	s A:	sn	Va	il Ti	`p /	75		r Hi	s Ala	Cys	leV 08
P	го	71	ור	Asp	P	ro	Asi 8	n Pi	го	Glr	n G	lu	Va		el L	.eu	Gli	J ASI	n Val	7hr 95	
н	is	Ph	e /	\sr	1 M	et 00	Trg	e Ly	ys .	Asr	n As	'n	Ме 10:		l G	lu	Glr	Me1	110	s Glu	Asp
Ī	l e	Ιt		er 15		eu	Trp	As	ip (	Gln	Se 12		Le	u Ly	s P	ro	Cys	Val		Leu	Thr
Pı	0	L e 13	u C 0	ys	V	al '	Thr	le	:u /	135	Су	s	Lys	s As	pV	al	Asn 140		Thr	Asn	Thr
T# 14	5	Ası	n A	sp	Se	er (	Slu	61 15	y 1	hr	Me	t	Gli	J Ar		l y 55	Glu	lle	Lys	Asn	Cys 160
Se	•	Pho	e A	sn	1 t	e 1	hr 165	Th	r S	er	11	e i	Arg	17		lυ	Val	Glr	lys	Gl u 175	Туг
AL	а	Lei	<b>9</b>	he	1 y 18	r L	ys.	Le	u A	SP	Va		Val 185		ı	le	Asp	Asn	Asn 190	Asn	Thr
Se	r	Tyi	- A:	rg 95	Le	u I	le	Sei	r C	ys	As <sub>1</sub>		ſhr	Sei	r Va	al	Ite	Thr 205	Gln	Ala	Cys
Pr	0 1	Lys 210	11	le	Se	r P	he	Glu	2 2	ro 15	H	e F	Pro	H	e Ni	İS	1 yr 220	Cys	Ala	Pro	Ala
22:	5							230	)						23	55				Lys	240
						24	45							250	)					Arg 255	
					260	)						2	65						270	Glu	
			27	5							280	)				•		285		1le	
	2	90							29	5						:	300			Gly	
305							3	10							31	5				lle	320
					•	32	5							330						Asn 335	
				3	40							34	15						350	Cys	-
			355	•						3	60							365		Thr	·
	37	0							37:	5						3	80			Ile	
385							39	90							399	•				Vol	400
Lys	Αl	<b>a</b>	let	Ty	<b>/</b> F	405	Pi	ro i	Pro	I	le	Ar		Gly 610	Glr	<b>1</b>	le i	Arg	Cys	Ser 415	Ser

G	ly I		35	le	rne A	rg P		40 40	iy Gi	y As	p Me	t Arg		) AST	1 Trp	
Aı		er 0	ilu l	.eu 1	yr L		yr Ly 55	ys Va	al Va	l Ly	s II		J Pro	Leu	Gly	
V 8		la P	rc 1	hr L		ia ty 70	ys Ar	rg Ar	g Va	t Va 47		n Arg	, Clu	Lys		
	-	NFOR	MAT I	ON F	OR S	_	NO:	25:		-7,	,					
		(i)	(A) (B) (C)	LEN TYP STR	CHAI GTH: E: ni ANDEI OLOG	1571 UCLEI ONESS	bas c ac	e pa id ngle	irs							
					TYPE	: DN	A (g	enom	ic)							
	(1	X)	(A) (B)	LOC	E/KEY ATION ER IN	ı: 1.	. 156									
	O	i) :	SEQU	ENCE	DESC	RIPT	ION:	SEQ	ID A	10:25	:					
Me	G GA t As	T GO	CA A'	IG A	NG AG /s Ar 5	A GG g Gl	G CT	C TG	C TG1 B Cys	Val	CTG Leu	CTG Leu	CTG Leu	TGT Cys 15	Gly	48
Ali	o Va	l Ph	e Va	l Se 20	r Pr	o Se	r Gli	n Glu 25		His	Ala	Arg	Phe 30	Arg	Arg	96
GG(	GC AL	a Ar	A AC	A GA ir Gl	A AA U Ly:	A TTI	G TGC J Trp 40	Val	ACA Thr	GTC Vai	TAT	TAT Tyr 45	GGG Gly	GTA Vel	CCT Pro	144
GT C Val	TG: Tr: 5:	D LY	G GA S Gl	A GC u Al	A ACI	C ACC	Thr	CTA	TTT Phe	TGT Cys	GCA Ala 60	Ser	GAT Asp	GCT ALa	AAA Lys	192
GCA Ala 65	Ty	GA As	T AC P Th	A GA r Gi	G GT/ U Vel 70	His	AAT Asn	GTT Vel	TGG Trp	GCC Ala 75	ACA Thr	CAT His	GCC	TGT Cys	GTA Val 80	240
CCC Pro	Thi	GAI AS	C CC	C AAI D ASI 8	n Pro	CAA GLm	GAA	GTA Val	GTA Val 90	TTG Leu	GTA Val	AAT Asn	GTG Val	ACA Thr 95	GAA Glu	288
AAT Asn	TT1 Phe	AAI ASI	0 AT	tTn	Lys	AAT	GAC Asp	ATG Met 105	GTA Val	GAA Glu	CAG Gln	ATG Met	CAT His 110	GAG Glu	GAT Asp	336
ATA	ATC	AGI Ser 115	Les	) Trp	GAT Asp	GLN	AGC Ser 120	CTA Leu	AAG Lys	CCA Pro	TGT Cys	GTA Val 125	AAA Lys	TTA Leu	ACC	384
CCA Pro	CTC Leu 130	Cys	GTT Vet	AGT Ser	TTA Leu	AAG Lys 135	TGC Cys	ACT Thr	GAT Asp	TTG Leu	GGG Gly 140	AAT Asn	GCT Ala	ACT Thr	AAT Asn	432
ACC Thr 145	AAT Asn	AGT Ser	AGT Ser	AAT	ACC Thr 150	AAT Asn	AGT Ser	AGT Ser	AGC Ser	GGG Gly 155	GAA Glu	ATG Het	ATG Het	ATG Met	GAG Glu 160	480
									AAT Asn							528

					10	65				17	0				175		
GG:	T AA	NG G	TG al	CA GL: 18	n L	AA G	AA T	AT GE yr A	CA TI La Pi 18	ne Ph	T TA	T AA/	A CTT s Leu	GAT ASP 190	He	ATA ile	576
Pro	1 A L	e A	AT SP 95	ASI	T GI n As	AT AC	or Ti	nr Se	C TA C Ty 00	AT AC	G TT(	S ACI	A AGT Ser 205	Cys	AAC Asn	ACC	624
TCA Ser	GT Va 21	ιI	TT Le	AC/ Thi	CA Gl	iG GC	C TO .a Cy 21	/S Pr	A AA	IG GT	A TCC L Ser	771 Phe 220	Glu	CCA Pro	ATT	CCC Pro	672
ATA 1 l e 225	Hi	T T	AT yr	TG1 Cys	GC Al	C CC a Pr 23	OAL	T GG	T TT y Ph	T GC: e Ala	G ATT B Ile 235	Leu	AAA I Lys	TGT Cys	AAT Asn	AAT Asn 240	720
AAG Lys	AC Th	G TI	ie .	AAT Asn	GG G1 24	y Ih	A GG r Gl	A CC y Pr	A TG O Cy	T ACA s The 250	A AAT r Asn D	GTC Val	AGC Ser	ACA Thr	GTA Val 255	CAA Gln	768
TCT Cys	AC:	A CA r Hi	\$	GGA Gly 260	110	T AG	g ac g Pr	A GT o Va	± :::T. I Va 26:	l Sei	A ACT	Gln	CTG Leu	CTG Leu 270	TTG Leu	AAT 'ASN	816
GGC Gly	AG1 Set	7 CT Le 27	u /	GCA Ala	GA	A GA	A GA	G GT. U Va 284	l Va	A ATT	AGA Arg	TET	GCC Ala 285	AAT Asn	TTC Phe	ACA Thr	864
Asp	Asr 290	AL	al	. <b>y</b> 5	Thr	· Ile	295 295	e Val	l Gli	let	AAC J Asn	Gln 300	Ser	Val	Glu	Ile	912
Asn 305	Cys	Th	r A	irg	Pro	310	t Asi	n Asr	n Thr	· Arg	Lys 315	Ser	ile	Arg	He	Gln 320	960
Arg	Gly	Pr	o G	ly	Arg 325	Ala	Phe	: Val	Thr	330		Lys	lle	Gly	Asn 335	Het	1008
Arg	Gln	Ala	3 H	15 40	Cys	Asn	ile	: Ser	345	Ala	Lys	Trp	Asn	Ala 350	Thr	Leu	1056
.ys	Gln	355	A	la	Ser	Lys	Leu	360	Glu	Gln	TTT Phe	Gly	Asn 365	Asn	Lys	Thr	1104
lle :	11e 370	Phe	. L)	ys (	Gin	Ser	Ser 375	Gly	Gly	Asp	Pro	Glu 380	He	Vəl	Thr	His	1152
er f 85	he	Asn	C	/s (	ily	Gly 390	Glu	Phe	Phe	Tyr	TGT Cys 395	Asn	Ser	Thr	Gln	Leu 400	1200
he A	sn	Ser	Th	r 1 4	rp .05	Phe	Asn	Ser	Thr	1rp 410	AGT Ser	Thr	Glu	Gly	Ser 415	Asn	1248
sn T	hr	Glu	G1 42	y S	er	Asp	Thr	lie	Thr 425	Leu	ECA Pro	Cys	Arg	1 le 430	Lys	Gln	1296
TT A he I	le.	AAC Asn 435	AT Me	G G	TG al	CAG Gln	GAA Glu	GTA Val 440	GGA Gly	AAA Lys	GCA Ala	Het	TAT Tyr 445	GCC Ala	CCT Pro	CCC Pro	1344

	Gly	CAA Gln									1392
		GGT Gly									1440
		GAT Asp									1488
		AAA Lys 500									1536
		GTG Val			TGA	GCG	G CC	GC			1571

- (2) INFORMATION FOR SEQ ID NO:26:
  - (i) SEQUENCE CHARACTERISTICS:
    - (A) LENGTH: 522 amino acids
    - (B) TYPE: amino acid
  - (D) TOPOLOGY: linear
  - (ii) MOLECULE TYPE: protein
  - (xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

Met Asp Ala Met Lys Arg Gly Leu Cys Cys Val Leu Leu Leu Cys Gly
1 5 10 15

Ala Val Phe Val Ser Pro Ser Gln Glu Ile His Ala Arg Phe Arg Arg 20 25 30 .

Gly Ala Arg Thr Glu Lys Leu Trp Val Thr Val Tyr Tyr Gly Val Pro 35 40 45

Val Trp Lys Glu Ala Thr Thr Thr Leu Phe Cys Ala Ser Asp Ala Lys 50 55 60

Ala Tyr Asp Thr Glu Val His Asn Val Trp Ala Thr His Ala Cys Val 65 70 75 80

Pro Thr Asp Pro Asn Pro Gln Glu Val Val Leu Val Asn Val Thr Glu 85 90 95

Asn Phe Asn Met Trp Lys Asn Asp Met Val Glu Gln Met His Glu Asp 100 105 110

Ile Ile Ser Leu Trp Asp Gin Ser Leu Lys Pro Cys Val Lys Leu Thr 115 120 125

Pro Leu Cys Val Ser Leu Lys Cys Thr Asp Leu Gly Asn Ala Thr Asn 130 135 140

Thr Asn Ser Ser Asn Thr Asn Ser Ser Ser Gly Glu Met Het Glu 145 150 155 160

Lys Gly Glu Ila Lys Asm Cys Ser Phe Asm Ile Ser Thr Ser Ile Arg 165 170 175

Gly Lys Val Gln Lys Glu Tyr Ala Phe Phe Tyr Lys Leu Asp Ite Ite 180 185 190

Pro Ile Asp Asn Asp Thr Thr Ser Tyr Thr Leu Thr Ser Cys Asn Thr

195 200 205

Ser Val Ile Thr Gln Ala Cys Pro Lys Val Ser Phe Glu Pro Ile Pro 210 215 220

Ite His Tyr Cys Ala Pro Ala Gly Phe Ala Ite Leu Lys Cys Asn Asn 225 230 235 240

Lys Thr Phe Asn Gly Thr Gly Pro Cys Thr Asn Val Ser Thr Val Gln 245 250 255

Cys Thr His Gly Ile Arg Pro Val Val Ser Thr Gln Leu Leu Asn 260 265 270

Gly Ser Leu Ala Glu Glu Glu Val Val lle Arg Ser Ala Asn Phe Thr 275 280 285

Asp Asn Ala Lys Thr Ile Ile Val Gin Leu Asn Gin Ser Val Giu Ile 290 295 300

Asn Cys Thr Arg Pro Asn Asn Asn Thr Arg Lys Ser Ile Arg Ile Gln 305 310 315 320

Arg Gly Pro Gly Arg Ala Phe Val Thr lie Gly Lys Ile Gly Asn Het 325 330 335

Arg Gln Ala His Cys Asn Ile Ser Arg Ala Lys Trp Asn Ala Thr Leu 340 345 350

Lys Gln 1le Ala Ser Lys Leu Arg Glu Gln Phe Gly Asn Asn Lys Thr 355 360 365

lie lie Phe Lys Gln Ser Ser Gly Gly Asp Pro Glu lie Val Thr His 370 380

Ser Phe Asn Cys Gly Gly Glu Phe Phe Tyr Cys Asn Ser Thr Gln Leu 385 390 395 400

Phe Asn Ser Thr Trp Phe Asn Ser Thr Trp Ser Thr Glu Gly Ser Asn 405 410 415

Asn Thr Glu Gly Ser Asp Thr Ile Thr Leu Pro Cys Arg Ile Lys Gln
420 425 430

Phe Ite Asn Met Val Gin Giu Val Gly Lys Ala Met Tyr Ala Pro Pro
435 440 445

Ite Ser Gly Gln Ite Arg Cys Ser Ser Asn Ite Thr Gly Leu Leu Leu 450 460

Thr Arg Asp Gly Gly Asn Asn Asn Gly Ser Glu Ile Phe Arg Pro 465 470 475 480

Gly Gly Gly Asp Met Arg Asp Asn Trp Arg Ser Glu Leu Tyr Lys Tyr 485 490 495

Lys Val Val Lys Ile Glu Pro Leu Gly Val Ala Pro Thr Lys Ala Lys 500 510

Arg Arg Val Val Gln Arg Glu Lys 515 520

- (2) INFORMATION FOR SEQ ID NO:27:
  - (i) SEQUENCE CHARACTERISTICS:
    - (A) LENGTH: 1532 base pairs (B) TYPE: nucleic acid
    - (C) STRANDEDNESS: single
    - (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA (genomic)

#### (ix) FEATURE:

- (A) NAME/KEY: CDS
  (B) LOCATION: 1..1522
- (D) OTHER INFORMATION:

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

	( ) 1	, 3	EGUE	MCE	UE 3C	K IP I	· UN :	35.0	ייטו	10:27	:					
	Asp			t Ly	G AGA S Arg					: Val						48
GCA Ala	GTC Val	Pho	C GT*	l Se	G CC( r Pro	Sei	C CAC	5 GA/ 1 Glu 25	ılle	CAT His	GCC	CGA Arg	TTC Phe 30	Arg	AGA Arg	96
Gly	GGC Gly	AG/ Arg 35	y Val	A GA	A AAG U Lys	i IT(	TGG Trp 40	Val	Thr	GTC Val	TAT	TAT Tyr 45	GGG Gły	GTA Val	CCT Pro	144
GTG Val	TGG Trp 50	Lys	GAA Glu	A GC/	A ACC	ACC Thr 55	Thr	CTA Leu	TŢŢ Phe	TGT Cys	GCA Ala 60	Ser	GAT Asp	GCT Alá	AAA Lys	192
GCA Ala 65	TAT Tyr	GAT	ACA Thr	GAC	GTA Val 70	His	AAT Asn	GTT Val	TGG Trp	GCC Ala 75	Thr	CAT	GCC Ala	TGT Cys	GTA Val 80	240
Pro	Thr	Asp	Pro	Asn 85		Gln	Glu	Val	Val 90	Leu	Glu	Asn	Val	Thr 95	Glu	288
His	Phe	Asn	100	Trp	Lys	Asn	Asn	Met 105	Val	Glu	Gln	Met	Gln 110	Glu	Asp	336
ATA I	ile	Ser 115	Leu	ТГР	Asp	Gln	Ser 120	Leu	Lys	Рго	Cys	Val 125	Lys	Leui	Thr	384
	130	Cys	Val	Thr	Leu	Asn 135	Cys	Lys	Asp	Val	Asn 140	Ala	Thr	Asn	Thr	432
ACT / Thr / 145	Asn .	Asp	Ser	Glu	Gly 150	Thr	Met	Glu	Arg	Gly 155	Glu	He	Lys	Asn	Cys 160	480
TCT 1 Ser F	he i	Asn	Ile	165	Thr	Ser	Ile	Arg	170	Glu	Val	Gln	Lys	6lu 175	Туг	528
GCT C Ala L	.eu f	Phe	190	Lys	Leu	Asp	Vat	Val 185	Pro	Ile	Asp	Asn	190	Asn	Thr	576
AGC T Ser T	yr A	195	Fen	He	Ser	Cys	Asp 200	Thr	Ser	Val	Ile	1hr 205	Gln	Ala	Cys	624
	ys I 10	le	Ser (	Phe	Glui	Pro 215	lle	Pro	lle	His	1yr 220	Cys	Ala	Pro	ALD	672
GGT T Gly PI 225	TT G he A	ice .	ATT (	.eu	AAG 1 Lys ( 230	rgt . Cys .	AAT Asn	GAT ASP	Lys	ACG Thr 235	TTC Phe	AAT Asn	GGA Gly	AAA Lys	GGA Gly 240	720

				sn V		SC AC				s th						768
ST. Va	A GT l Va	A TO	er It	CT CA nr G 50	MA CT	G CT	G CT.	A AA u Asi 26	n Gl	C AG	T CTA	A GCA J Ala	GAA Glu 270	Glu	GAG Glu	816
			e Ar			C AA1		e Thi					Thr			864
		n Le				T GTA r Val 295	GL					Arg				912
	ı Ih					A CAT e His O					Arg					960
ACA Thr	GG	A GA y Gl	TA A JI u	A AT e Il 32	e Gl	A GAT y Asp	ATA	AGA Arg	CAA Gln 330	ALa	CAT His	TGT Cys	AAC Asn	ATT Ile 335	AGT Ser	1008
				p As		C ACT			Gln							1056
GAA Glu	CAJ Glr	1 TT n Ph	e Gl	G AA' u Asi	T AA/	ACA Thr	ATA Ile 360	Val	Phe	AAT Asn	CAC	TCC Ser 365	TCA Ser	GGA Gly	GGG Gly	1104
GAC Asp	CCA Pro 370	Gli	A ATT	r GT/ e Vai	A ATG	CAC His 375	AGT Ser	TTT Phe	AAT Asn	TGT Cys	GGA Gly 380	GGA Gly	GAA Glu	TTT Phe	TTC Phe	1152
TAC Tyr 385	Cys	AA1 Asr	T TCA	ACA Thr	Gln 390	CTG Leu	TTT Phe	AAT Asn	AGT Ser	ACT Thr 395	TGG Trp	AAT Asn	AAT Asn	AAT Asn	ACT Thr 400	1200
GAA Glu	GGG Gly	Ser	AA1 Asr	AAC Asn 405	Thr	GAA Glu	6GA 61y	AAT Asn	ACT Thr 410	ATC Ile	ACA Thr	CTC Leu	CCA Pro	TGC Cys 415	AGA Arg	1248
ATA Ile	AAA Lys	Glr	420	He	AAC Asn	ATG Met	GTG Val	CAG Gln 425	GAA Glu	GTA Val	GGA Gly	AAA Lys	GCA Ala 430	ATG Met	TAT Tyr	1296
GCC Ala	CCT Pro	CCC Pro 435	lle	AGA Arg	GGA Gly	CAA Gln	ATT Ile 440	AGA Arg	TGT Cys	TCA Ser	TCA Ser	AAT Asn 445	ATT Ile	ACA Thr	ela	1344
CTG Leu	CTA Leu 450	TTA Leu	ACA Thr	AGA Arg	GAT Asp	GGT Gly 455	G <b>GT</b> Gly	ATT	AAT Asn	GAG Glu	AAT Asn 460	GGG Gly	ACÇ Thr	GAG Glu	ATC Ile	1392
TTC Phe 465	AGA Arg	CCT Pro	GGA Gly	GGA Gly	GGA Gly 470	GAT . Asp i	ATG Met	AGG Arg	Asp	AAT Asn 475	TGG Trp	AGA Arg	AGT Ser	GAA Glu	TTA Leu 480	1440
TAT	AAA Lys	TAT Tyr	AAA Lys	GTA Val 485	GTA Val	AAA I	ATT Ile	Glu	CCA Pro 490	TTA Leu	GGA Gly	GTA Val	GCA Ala	CCC Pro 495	ACC Thr	1488
						GTG ( Val (	iin .				T GA	ccec	ccGC	:		1532

- (2) INFORMATION FOR SEQ ID NO:28:
  - (i) SEQUENCE CHARACTERISTICS:
    - (A) LENGTH: 507 amino acids
    - (B) TYPE: amino acid
    - (D) TOPOLOGY: Linear
  - (ii) MOLECULE TYPE: protein
  - (x1) SEQUENCE DESCRIPTION: SEQ ID NO:28:
- Met Asp Ala Met Lys Arg Gly Leu Cys Cys Val Leu Leu Leu Cys Gly
  1 5 10 15
- Ata Val Phe Val Ser Pro Ser Gln Glu Ile His Ala Arg Phe Arg Arg 20 25 30
- Gly Gly Arg Val Glu Lys Leu Trp Val Thr Val Tyr Tyr Gly Val Pro 35 40 45
- Val Trp Lys Glu Ala Thr Thr Thr Leu Phe Cys Ala Ser Asp Ala Lys
  50 60
- Ala Tyr Asp Thr Glu Val His Asn val Trp Ala Thr His Ala Cys Val 65 70 75 80
- Pro Thr Asp Pro Asn Pro Gln Glu Val Val Leu Glu Asn Val Thr Glu 85 90 95
- His Phe Asn Met Trp Lys Asn Asn Met Val Glu Gln Met Gln Glu Asp 100 105 110
- Ite Ite Ser Leu Trp Asp Gln Ser Leu Lys Pro Cys Val Lys Leu Thr 115 120 125
- Pro Leu Cys Val Thr Leu Asn Cys Lys Asp Val Asn Ala Thr Asn Thr 130 135 140
- Thr Asn Asp Ser Glu Gly Thr Het Glu Arg Gly Glu Ile Lys Asn Cys 145 155 160
- Ser Phe Asn Ile Thr Thr Ser Ile Arg Asp Glu Val Gln Lys Glu Tyr 165 170 175
- Ala Leu Phe Tyr Lys Leu Asp Val Val Pro Ile Asp Asn Asn Asn Thr 180 185 190
- Ser Tyr Arg Leu Ile Ser Cys Asp Thr Ser Val Ile Thr Gln Ala Cys
  195 200 205
- Pro Lys Ile Ser Phe Glu Pro Ile Pro Ile His Tyr Cys Ala Pro Aia 210 215 220
- Gly Phe Ala Ile Leu Lys Cys Asn Asp Lys Thr Phe Asn Gly Lys Gly 225 235 240
- Pro Cys Lys Asn Val Ser Thr Val Gln Cys Thr His Gly Ile Arg Pro 245 250 255
- Val Val Ser Thr Gln Leu Leu Leu Asn Gly Ser Leu Ala Glu Glu Glu Glu 260 265 270
- Val Val Ile Arg Ser Asp Asn Phe Thr Asn Asn Ala Lys Thr Ile Ile 275 280 285
- Val Gin Leu Lys Glu Ser Val Glu Ile Asn Cys Thr Arg Pro Asn Asn 290 295 300
- Asn Thr Arg Lys Ser Ile His Ile Gly Pro Gly Arg Ala Phe Tyr Thr

305 The Bly Glu Ite Ite Gly Asp Ite Arg Glo Ata His Cys Ash Ite Ser arg Ala Lys Tro Asn Asp Thr Leu Lys Gln Ile Val Ile Lys Leu Arg Glu din Phe Glu Asn Lys inn lie val Phe Asn His Ser Ser Gly Gly 355 360 365 Asp Pro Glu Ite Val Met His Ser Phe Asn Cys Gly Gly Glu Phe Phe Tyr Cys Asn Ser Thr Gln Leu Phe Asn Ser Thr Trp Asn Asn Asn Thr 385 390 395 400 Glu Gly Ser Asn Asn Thr Glu Gly Asn Thr Ile Thr Leu Pro Cys Arg Ile Lys Gln Ile Ile Asn Met Val Gln Glu Val Gly Lys Ala Met Tyr 420 425 430 Ala Pro Pro Ile Arg Gly Gin Ile Arg Cys Ser Ser Asn Ile Thr Gly
435 440 445 Leu Leu Leu Thr Arg Asp Gly Gly Ile Asn Glu Asn Gly Thr Glu Ile 450 455 460 Phe Arg Pro Gly Gly Gly Asp Met Arg Asp Asn Trp Arg Ser Glu Leu 465 470 475 480 Tyr Lys Tyr Lys Val Val Lys Ile Glu Pro Leu Gly Val Ala Pro Thr 485 490 495 Lys Ala Lys Arg Arg Val Val Gin Arg Glu Lys

- (2) INFORMATION FOR SEQ ID NO:29:
  - (i) SEQUENCE CHARACTERISTICS:
    - (A) LENGTH: 15 amino acids
    - (B) TYPE: amino acid
    - (C) STRANDEDNESS: single (D) TOPOLOGY: linear
  - (ii) MOLECULE TYPE: protein
  - (xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:

Ala Lys Arg Arg Val Val Gin Arg Glu Lys Arg

### What is claimed is:

- A recombinant nucleic acid molecule which encodes a mutant HIV-1 gp120 envelope glycoprotein comprising a V3 loop deletion and a C4 domain<sub>(W->X)</sub> point mutation, wherein X is an amino acid residue other than tryptophan.
- The recombinant nucleic acid molecule of claim 1,
   wherein X is a valine residue.
  - 3. The recombinant nucleic acid molecule of claim 1, wherein the nucleic acid molecule is a DNA molecule.
- 15 4. The recombinant nucleic acid molecule of claim 3, wherein the DNA molecule is a plasmid.
- 5. The recombinant nucleic acid molecule of claim 4, wherein the plasmid comprises the sequence of the plasmid designated PPI4-tPA.
  - 6. The recombinant nucleic acid molecule of claim 1, wherein the C4 domain is an HIV-1<sub>IAI</sub> gp120 envelope glycoprotein C4 domain.

- 7. The recombinant nucleic acid molecule of claim 6, wherein the mutant HIV-1 gp120 envelope glycoprotein is a mutant HIV-1 gp120 envelope glycoprotein.
- 30 8. The recombinant nucleic acid molecule of claim 1, wherein the C4 domain is an HIV-1 $_{\rm IR-FL}$  gp120 envelope glycoprotein C4 domain.
  - 9. The recombinant nucleic acid molecule of claim 8,

WO 94/22477

wherein the mutant HIV-1 gp120 envelope glycoprotein is a mutant HIV-1 $_{\rm JR-FL}$  gp120 envelope glycoprotein.

- 10. The mutant HIV-1 gp120 envelope glycoprotein encoded by the recombinant nucleic acid molecule of claim 1.
  - 11. A vaccine which comprises a therapeutically effective amount of the mutant HIV-1 gp120 envelope glycoprotein of claim 10, and an adjuvant.

10

12. A method of treating an HIV-1-infected subject, which comprises immunizing the HIV-1-infected subject with the vaccine of claim 11, thereby treating the HIV-1-infected subject.

15

- 13. A vaccine which comprises a prophylactically effective amount of the mutant HIV-1 gp120 envelope glycoprotein of claim 10, and an adjuvant.
- 20 14. A method of reducing the likelihood of an HIV-1-exposed subject's becoming infected with HIV-1, which comprises immunizing the HIV-1-exposed subject with the vaccine of claim 13, thereby reducing the likelihood of the HIV-1-exposed subject's becoming infected with HIV-1.

25

- 15. A method of reducing the likelihood of a non-HIV-1-exposed subject's becoming infected with HIV-1, which comprises immunizing the non-HIV-1-exposed subject with the vaccine of claim 13, thereby reducing the likelihood of the non-HIV-1-exposed subject's becoming infected with HIV-1.
- 16. A method of obtaining partially purified antibodies which specifically bind to the CD4-binding domain of

- HIV-1 gp120 envelope glycoprotein, which method comprises (a) immunizing a non-HIV-1-exposed subject with the vaccine of claim 13, (b) recovering from the immunized subject serum comprising said antibodies, and (c) partially purifying said antibodies, thereby obtaining partially purified antibodies which specifically bind to the CD4-binding domain of HIV-1 gp120 envelope glycoprotein.
- 10 17. The method of claim 16, wherein the subject is a human.
  - 18. The partially purified antibodies produced by the method of claim 16.
- 19. A pharmaceutical composition, which comprises a therapeutically effective amount of the partially purified antibodies of claim 18, and a pharmaceutically acceptable carrier.
- 20 20. A method of treating an HIV-1-infected subject, which comprises administering to the subject a dose of the pharmaceutical composition of claim 19 effective to reduce the population of HIV-1-infected cells in the HIV-1-infected subject, thereby treating the HIV-1-infected subject.
- 21. A method of treating an HIV-1-infected subject, which comprises administering to the subject a dose of the pharmaceutical composition of claim 19 effective to reduce the population of HIV-1 in the HIV-1-infected subject, thereby treating the HIV-1-infected subject.
- 22. A composition which comprises a prophylactically effective amount of the partially purified antibodies of claim 18, and a pharmaceutically acceptable carrier.

WO 94/22477 PCT/US94/03282

90

- 23. A method of reducing the likelihood of an HIV-1-exposed subject's becoming infected with HIV-1, which comprises administering to the HIV-1-exposed subject a dose of the composition of claim 22 effective to reduce the population of HIV-1 in the HIV-1-exposed subject, thereby reducing the likelihood of the subject's becoming infected with HIV-1.
- 10 24. The method of claim 23, wherein the subject is a medical practitioner.
  - 25. The method of claim 23, wherein the subject is a newborn infant.

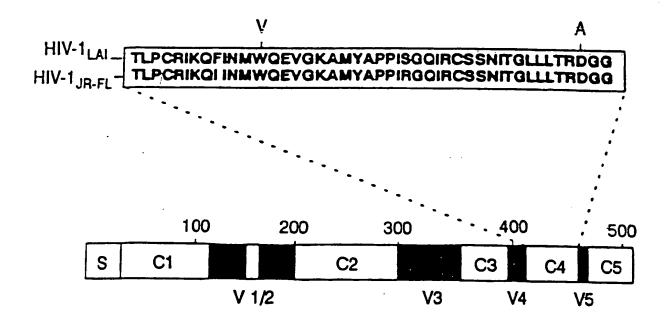
A method of reducing the likelihood of a non-HIV-1-

15

26.

- exposed subject's becoming infected with HIV-1 as a result of exposure thereto during an incident wherein there is an increased risk of exposure to HIV-1, which comprises administering to the subject immediately prior to the incident a dose of the composition of claim 22 effective to reduce the population of HIV-1 to which the subject is exposed during the incident, thereby reducing the likelihood of the subject's becoming infected with HIV-1.
  - 27. The method of claim 26, wherein the subject is a medical practitioner.

FIGURE 1



## FIGURE 2

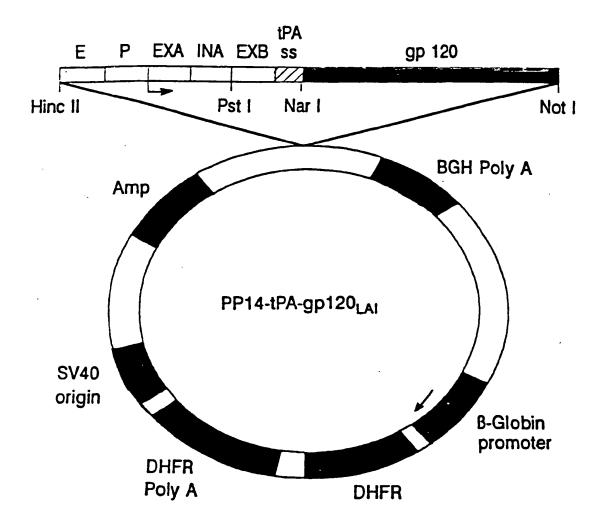


FIGURE 3A

FIGURE 3A FIGURE 3D FIGURE 3D FIGURE 3E FIGURE 3E ttgacattgattattgactagttattaatagtaatcaattacggggtcattagttcatagcccatatatgga gttocgogttacataacttacggtaaatggcocgoctggctgaccgcccaacgaccccccgcccattgacgtc aataatgacgtatgttcccatagtaacgccaatagggactttccattgacgtcaatgggtggactatttacg gtanactgeceacttggeagtacateaagtgtateatatgeeaagtaegeeeeetattgaegteaatgaegg attagtcatcgctattaccatggtgatgcggttttggcagtacatcaatgggcgtggatagcggtttgactc taaatggcccgcctggcattatgcccagtacatgaccttatgggactttcchacttggcagtacatctacgt acggggatttccaagtctccaccccattgacgtcaatgggaytttgtttg@caccaaaatcaacgggactt 289 145 217 361 433

SUBSTITUTE SHEET (RULE 26)

HincII

# FIGURE 3B

505

SUBSTITUTE

4/42 tccaaaatgtcgtaacaactccgccccattgacgcaaatgggcggtaggcgtgcggtggggaggtctatat aagcagagetegtttagtgaaccGTCAGATCGCCTGGAGGCCCATCCACGCTGTTTGACCTCCATAGAAG **ACACCGGGACCGATCCAGCCTCCGCGGCCGGGAACGGTGCATTGGAACGCGGATTCCCCGTGCCAAGA**GTGA ggccaacaccccgtcctagataggtgatggtatagcttagcctataggtgtgggttattgaccattattgac ctccacgcgaatctcgggtacgtgttccggacatgggctcttctccggtagcggcggagctccacatccgag cctgtcccatgcccatgcctccagggctcatggtcgctcggcagctccttgctcctaacagtggaggccag cactcccctattggtgacgatactttccattactaatccataacatggccgctctttgccacaactatctct attggctatatgccaatactctgtccttcagagactgacacggactctgtnttttacaggatggggtccca tttattatttacaaattcacatatacaacaacgccgtcccccgtgcccgcagttttattaacatgcgggat acttaggcacaggacaatgcccaccaccagtgtgcgcacaaggccgtggcggtagggtatgtgtctga **▼** Transcription Start Intron A **Exon A** 577 649 793 1009 721 865 937 1081 1153 1225

SHEET (RULE 26)

## IGURE 3C

1297	aaatgageteggagattgggetegeaeegetgaegeatggaagaettaaggeagegeggeagaagaagatge
1369	aggcagctgagttgttattctgtagagttggaggtaactcccgttgcggtgctgttaacggtggagggca
1441	gegegecaecagaeata
1513	PstI Exon B tcctttccatgggtctttctgcagTCACCGTCCTTGACACGATGGATGCAATGAAGAGAGGGCTCTGCTGT M D A M K R G L C C
1585	Nari GTGCTGCTGTGTGTGGAGCAGTCTTCGCCCAGCCAGGAAATCCATGCCCGATTCAGAAGAGGCGCC V L L L C G A V F V 8 P 8 Q E I H A R F R R G A
1657	AGAACAGAAAATTGTGGGTCACAGTCTATTATGGGGTACCTGTGTGGAAGGAA
1729	▲ Signal cleavage rerecarcasargetalacasasasasasatacasasasatasatasatasatasasasas
29	CASDAKAYDTEVHNVNATHACVPT
1801	GACCCCAACCCACAAGAAGTAGTATTGGTAAATGTGACAGAAAATTTTAACATGTGGAAAAATGACATGGTA D P N P Q E V V L V N V T E N F N M W K N D M V

SUBSTITUTE SHEET (RULE 26)

# FIGURE 3D

	1873	S&	CAG	ATG	CAT	GAG	GAT	ATA	ATC	.AG1	TIA	756	GAT	3	Section	CIA	AAG	CC	TGT	GTN	. A	Z T Z		ۯ	) L
	101	កោ	0	Σ	×	m.	Ω	H	H	တ	H	EQMHEDIISLWDOSLKPC.KLTPL	۵	0	S	H	×	۵,	ပ	V.	×		) H	מ ל	ָרָ בְּיִבְּי
7	1945	TGT	GII	AGT	TTA	<b>3</b> 46	1100 1100	ACI	GAI	TIC	9	AAT	GCT	ACT	AAT	ACC.	TAAT	AGT	AGT	AAT		FAA	T.C.	בטע	ر ا
<b>~</b>	131	ပ	>	လ	<b></b>	×	U	H	Ω		G	CVSLKCTDLGNATNTSSNTNSSS	~	H	z	E+	z	တ	တ	z	<b>H</b>	z	် လ	S	S
	2017	9	8	ATG	ATG	ATG		3	53	GAG	ATA	***************************************	<b>X</b> C	ည်	ICI	TTC	ZAT	ATC	AGC	2		ATA	AGA	ניט	AAG
	155	ပ	ப	I	Σ	Σ	M	×	G	M	H	GEMMMEKGEIKNCSFNISTSIRGK	Z	ပ	တ	Em.	Z	H	တ	€→	ဟ	н	æ	ט	×
~	2089	GTG	SAS	₹	8	TAT	Ş	TTI	TTI	TAI	3	CTT	GAT	ATA	ATA	Ş	ATA	GAT	Z.	GAT	ACT	ပ္သ	ည	TAT	ACG
	179	V Q K E Y A F F Y K L D I I P I D N D T T S Y T	0	×	M	<b>&gt;</b>	~	<u>fa</u>	<b>[24</b>	<b>&gt;</b>	×	H	Ω	H	<b>H</b>	<u>Ω</u> ,	H	۵	z	Ω	F	£+	တ	<b>&gt;</b>	H
~	2161	TTG	ঠু	AGT	TGT	<b>X</b> C	<b>S</b> CC	ភ្ជ	GIC	ATI	Ş	SSS	ပ္ပ	TGT	ర్ట	<b>3</b> 46	GTA	TCC	TTT(	GAG	S	ATT	ပ္ပ	ATA	CAT
	203	LTSCNTSVITOACPKVSFEPIH	H	တ	ပ	Z	Ð	တ	>	<b>H</b>	H	0	<b>«</b>	ပ	۵,	×	>	S	<u>[a,</u>	ഥ	۵.	H	Q.	H	<b>=</b>
7	2233	TAT	TGT	႘	TATTGTGCCCCGGCI	SCI	561	TTT	ပ္ပ	ATT	CTA	rggttttgcgattctaaaatgtaataataagacgttcaatggaacaggaccatgtaca	TGT	AAT	MAT	MAG	ACG	TTC	AAT	GGA	Ž	GGA	ర్ట	TGI	ACA
	227	<b>&gt;</b> +	ပ	×	Y C A P	K	ၒ	Ĺ	4	H	H	×	ပ	Z	Z	×	H	(La)	z	ပ	H	ပ	۵.	ပ	H
. ~	2305	AAT.	GTC	<b>A</b> G	ACA	GTA	3	TGT	Ş	CAT	₹9	ATT	<b>8</b> 66	Ş	GTA	GTA	Ş	AC.T.	Š	CT	CTG	ITG	MAT	ပ္ပ	AGT
l	251	N V S T V Q C T H G I R P V V S T Q L L N G S	>	တ	۲	>	O	ပ	H	x	U	H	Œ	<u>م</u>	>	>	S	H	O	H	1	.a	Z	ပ	S
C	7477	CTAGCAGIAGAAGAGAGTAATTAGATCTGCCAATTTCACAGACAATGCTAAAACCATAATAGTACAGCTG	į		<b>3</b>	GAG	GTA	GTA	ATT	AGA	TCT	ပ္ပ	<b>₹</b>	TTC	Ş	S	AAT	SCT	3	D D	MTA	ATA(	GTA	CAG	CTG
1	275	J	4	ம	ш	ப	>	>	H	<b>c</b> .	ഗ	~	z	ſ.	H	۵	Z	~	×	H	H	H	>	Ci	H

6/42

SUBSTITUTE SHEET (RULE 26)

FIGURE 3E

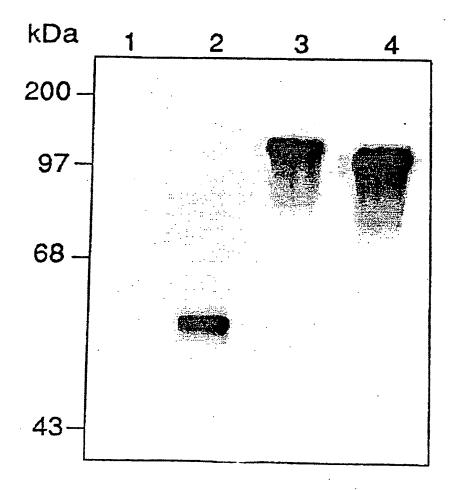
<b>გ</b> ენ	9C <b>4</b>	ATC I	TAC	ACT T	₹ ×	AGA R	AGA R
<b>A</b> QQ	AGA R	AAATGGAATGCCACTTTAAAACAGATAGCTAGCAAATTAAGAGAACAATTTGGAAATAATAAAACAATAATCAATAATCAATAATCAATAATCAATAAT	TIC	t <b>gtaattcaacacaactgtttaat</b> agtacttggtttaatagtacttggagtactgaagggtcaaataacat c n s t q l f n s t w f n s t w s t e g s n n t	GAAGGAAGTGACACACACCCCATGCAGAATAAAACAATTTATAAACATGTGGCAGGAAGTAGGAAAA E G S D T I T L P C R I K Q F I N M W Q E V G K	GCAATGTATGCCCCTCCCATCAGGGGCACAAATTAGATGTTCATCAAATATTACAGGGCTGCTATTAACAAGA A M Y A P P I S G Q I R C S S N I T G L L L T R	GATGGTGGTAATAACAACAATGGGTCCGAGATCTTCAGACCTGGAGGAGGAGATATGAGGGACAATTGGAGA D G G N N N G S E I F R P G G G D M R D N W R
5	AGT S	ACA T	TTT	AAT	GTA V	TTA	AA
ATC	ATT	₹ ×	E E	TCA s	GAA	CTA	GAC
CGT	Z	Z z	ეტ ე	<del>ပို့</del> ဗ	CAG	CTG	<b>A</b> 66
ATC	TGT C	AAT N	<b>წ</b> ა	GAA E	77GG ₩	999	ATG
AGT S	CAT	့ ဗို့	រត្តិរ ភ	ACT H	ATG	ACA T	SAT
<b>₹</b> ×	ð Z	TTT	AAT	AGT S	AAC	ATT I	<b>წ</b>
AGA R	30	<b>₹</b> ∘	TTT	7GG ¥	ATA I	AAT	<b>წ</b> ე
A F	AGA	GAA	AGT S	ACT T	TTT	TCA S	ტე ე
AAT	ATG M	Ag «	ည္	AGT	<b>₹</b> ○	TCA S	CCT
Z Z	AAT	TT.	<b>A</b> CG →	AAT	<b>₹</b> ×	រក្សា ភ	AGA
Z Z	ဗွ် ပ	₹ ×	GTA >	ttt F	ATA I	AGA R	TIC
ပ္ပ	ATA	ည္မွ	ATT	નુ જ	AGA R	ATT	ATC
AGA R	<b>₹</b> ×	ည် •	<b>Ж</b> ш	ACT	ပ္ခ် ပ	30	GAG E
A F	(ပို့ ပ	ATA I	ပ္လို ေ	AGT	ည်မှ	<b>္ပြီ</b> ပ	TCC s
ក្សិ	ATA	30	ည္သို့ သ	Z	CTO 1	<b>A</b> GC 8	ပ္ပ် ပ
A	N F	<b>₹</b> ×	წე	111	S F	ATC	Z Z
ATT	GTT V	TTA	<b>წ</b>	D 1	ATC	ပ္ပ် ရ	<b>2</b> 2
E E	TTT	ACT	ညီ လ	30	N H	ر 1	Z
GTA V	ર્ધું ∢	ပ္ပဲ <	ာ်င် န	A P	GAC	ეე. •	Z Z
TCT S	AGA R	X	30	ညီ လ	AGT	TAT Y	င်ဌာ ရေ
30	ပ္တ	9 25 <b>≥</b>	×	Z Z	<b>(</b> 000)	ATG	G
AACCAATCTGTAGAAATTAATTGTACAAGACCCAACAACAATACAAGAAAAAGTATCCGTATCCAGAGGGGA N Q S V E I N C T R P N N T R K S I R I Q R G	CCAGGGAGACATTGTTACAATAGGAAATATGAGACAAGCACATTGTAACATTAGTAGAGCA PGRAFVTIGTACATTAGTAGAAATAGGAAATATGAGACAAGCACATTGTAACATTAGTAGAGCA	<b>X</b> ×	TTTAAGCAATCCTCAGGGGGGCCCAGAATTGTAACGCACAGTTTTAATTGTGGGGGGGAATTTTTCTAC F K Q S S G G D P E I V T H S F N C G G E F F Y	ភពិ	S a	ეე ▼	GAT
2449 299.	2521 323	2593 347	2665	2737 395	2809	2881	2953
	SHE	RSTITLI	TF SHF	FT (RI	II E 261	<b>L</b>	

FIGURE 3F

Œ GTGGTGCAGAGAAAAATGAGCGCCCC NotI æ O 3025 3097

SUBSTITUTE SHEET (RULE 26)

FIGURE 4



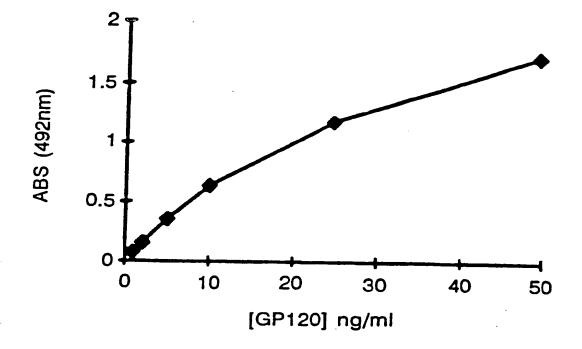
WO 94/22477 PCT/US94/03282

FIGURE 5A

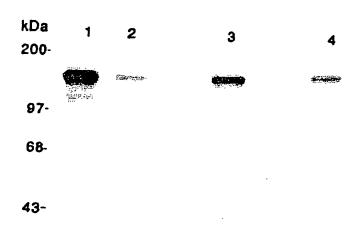
Stable CHO	[gp120]
clone	(ng/ml)
5	6
6	14
9	123
10	4
12	18
13	18

WO 94/22477 PCT/US94/03282

FIGURE 5B



## FIGURE 6



29-

FIGURE 7A

FIGURE 7C FIGURE 7A FIGURE 7B

ATGGATGCAATGAAGAGA

≥ ⋖ Ω

O ဟ ۵. ⋖ U U NarI 19

CATGCCCGATTCAGAAGAGGCGCAGAGTAGAAAAGTTGTGGGTCACAGTCTATTATGGG ഥ > **5** Œ Œ Œ ⋖

GTACCTGTGTGGAAAGAAGCAACCACCACTCTATTTTGTGCATCAGATGCTAAAGCATAT 0 S cleavage ⋖ U Signal 4 M ×

GATACAGAGGTACATAATGTTTGGGCCACACATGCCTGTGTACCCACAGACCCCAACCCA Ω ρ, > U < × Z ×

Caagaagtagtattggaaatgtaacagaacattttaacatgtggaaaaataacatggta 259

gaacagatgcaggagatataatcagtttatgggatcaaagcctaaagccatgtgtaaa Z E Z <u>L</u> M > Z M 319 87

**TTAACCCCACTCTGTGTTACTTTAAATTGCAAGGATGTGAATGCTACTAATACCACTAAT** ပ Ä S 0 Δ H S H H Ω P 0 E 0 379 107

GATAGCGAGGGAACGATGGAGAGAGGAGAATAAAAAACTGCTCTTTCAATATACACCACA တ ы G æ 439 147

Ľ

G

Ω

×

ပ

127

JR-FL

SUBSTITUTE SHEET (RULE 26)

199

67

139

## URE 7B

499 167	AGCATAAGAGATGCGCAGAAAGAATATGCTCTTTTTTATAAACTTGATGTAGTACCA S I R D E V Q K E Y A L F Y K L D V V P	<b>3</b>	<b>2 2 2 3 3 3 3 3 3 3 3 3 3</b>	SAT	S a	95 55 5	30	<b>₹</b> ×	<b>E</b> 2	TAT X	<b>₩</b>	CTT 1	FIT	TAT	Ž×	CTŢ	GAT	STA	GTA(	<b>6</b>
559 187	ATAG	ATA	Z Z	Z Z	AAT	D F	သူလ	TAT	<b>7</b> 66	TTG	SATA	<b>AG1</b> S	161 C	ည်ရ	₹CC ₽	TC. S	GTC.	ATT	Ž +	SAG O
619 207	GCCIGICCAAAGATAICCTIIGAGCCAATICCCAIACAITAIIGIGCCCCGGCIGGIIII A C P K I S F E P I P I H Y C A P A G F	GIC	200	3 ×	ATA I	S S	TIT	S E	ပ္သိ 🏊	17	ည	ATA:	EAT H	TAT	<b>1</b> 61	ပ္ပ	ပ္ပံွ	SCT A	ပည် ပည်	TIT
679 227	GCGA A	IIC	Ĭ, J	Z ×	ည်	X	GAT	×	<b>₩</b>	ITT	ZA1	<b>ြ</b> ပ်	<b>3</b> ×	<b>6</b> 69	<b>5</b> 0	TGT	A ×	AAT	GTC V	နှင့်င လ
739	ACAG	TAC	<b>3</b> 0	5 F	Ž H	E E	<b>်</b> ပြွ	ATT	<b>8</b>	ပ္ပ	\GT2	\GT2	NTCA S	<b>5</b> 64	<b>₹</b> 0	- 51.73 - 51.73	CIG	CTA	AAT	ပ္ပ ဖ
9 ~	AGTC	TAC	Ž Z	E SA	E B	E E	CTA V	CTA V	LTA I	AG R	ATCT S	CGAC	AAT N	TTC	ACG T	N Z	AAT	GCT A	¥ ×	ACC
859 287	ATAA	I	ST X	80	CIG	X	E E	S	GT3	IGA.	277	Z Z	7. 1.01	- ACA	ATAGTACAGCTGAAAATCTGTAGAAATTAATTGTACAAGACCCAACAACAATACA I V Q L K E S V E I N C T R P N N N T	ည	AAC N	N N	AAT	A CA
919	AGAA	Ž	NGTJ S	ATA	EAT H	ATA	ည်ပ	ည်ရှ	တ္တိမ	SAG!	\QC\ \\	(TT) F	rtat Y	ACT T	ACA	်ပ္လို့ ပ	S a	ATA	ATA I	AAAAGTATACATATAGGACCAGGGAGCATTTTATACTACAGGAGAAATAATAGGA K S I H I G P G R A F Y T T G E I I G
979 327	GATA D	Z H	NGA(	80	ర్ట్ ∢	CAT	ည်	Z Z	:ATT	rag1 S	rag.	<b>1</b> 00 <b>⋖</b>	A X	\TGG ¥	N	GAC	ACT	TTA	₹ ×	ATAAGACAAGCACATTGTAACATTAGTAGAGCAAAATGGAATGACACTTTAAAACAG I R Q A H C N I S R A K W N D T L K Q

14/42

SUBSTITUTE SHEET (RULE 26)

FIGURE 7C

1039	ATA	\GT1	ATA	3×	LTT.	202	S a	30	TTT F	ATAGTTATAAAATTAAGAGAACAATTGAGAATAAAACAATAGTCTTTAATCACTCCTCA I V I K L R E Q F E N K T I V F N H S S	Z Z	*	ACA T	ATA	GTC >	TIT	AAT	CAC	TCC လ	TC S	æ
1099	96 <b>A</b> (	999	SGAC D	ည်း	SE E	E H	rGT.	Æ Æ	S. H	GGGGACCCAGAAATTGTAATGTAATTGTGGAGAGAATTTTTCTACTGT G D P E I V M H S F N C G G E F F Y C	TII	X Z	TGT C	<b>ဗိ</b> ဗ္ဗ	<b>წ</b> ე	<b></b>	TTT	TIC	ZI ≯	မြို့ပ	H
1159	AA'S N	ည်လ	Ž t	30	CT	TT:	Z	AG1 S	1. ACT	AA'ITCAACACAACTGTTTAATAGTACTTGGAATAATAATAATGGGTCAAATAACACT N S T Q L F N S T W N N T E G S N N T	N	AAT	AAT	ACT	GAA	့ ဗိဗ္ဗ	S	A Z	Z Z	Ş F	H
1219	GA.	(၁)	EAS N	13C1	ATC I	14C)	CTC	ប្តី 🕰	ည်ပ	GAAGGAAATACTATCACACTCCCATGCAGAATAAAACAAATTATAAACATGTGGCAGGAA E G N T I T L P C R I K Q I I N M W Q E	ATA	\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\	80	ATT	ATA	ZZ	ATG	7 ₹	OA O	E G	4
1279	GTA V	နှင့် ၁	₹×	Çy <b>∢</b>	X X	STA1	ည္မ	) (1)	ည	GTAGGAAAAGCAATGTATGCCCCTCCCATCAGAGGACAAATTAGATGTTCATCAAATATT V G K A M Y A P P I R G Q I R C S S N I	AGA	<b>်</b> ပ္ပြဲ ပ	80	ATT	<b>A</b> G <b>A</b>	igr C	TCA	ည်းလ	ZAZ	TA. I	H
1339	ACA T	တ္တဗ	CTG 1	i E	TT a	N P	<b>10 2 2</b>	GAT	ည်ဝ	ACAGGGCTGCTATTAACAAGAGATGGTGGTATTAATGAGAATGGGACCGAGATCTTCAGA T G L L L T R D G G I N E N G T E I F R	ATT	AAT	GAG	MAT	တ္တ ဗ	_ <del>}</del> €	GAG	ATC	TIC	9 % **	<b>«</b>
1399	CCT	ဦ် ပ	<b>(</b> 5)	<b>(</b> 5)	GA1 D	ratg M	<b>1</b> 866	<b>1</b> 080	X	CCTGGAGGAGAGATATGAGGGACAATTGGAGAAGTGAATTATATAAATATAAAGTAGTA P G G G D M R D N W R S E L Y K Y K V V	AGA R	AGT	E G	LL	TAT	<b>₹</b> ×	TAT.	Ž×	GT?	ST.	<
1459	× A	ATT	<b>₹</b> 9	<b>(</b> 2)	TT	<b>1</b> 5000	GTA	ઈું ∢	ည်မှ	ATTGAACCATTAGGAGTAGCACCCACCAAGGCAAAGAGAAGAGTGGTGCAAAGAGAA I E P L G V A P T K A K R N V V Q R E	AAG	ည္မွ	<b>AA</b> G	AGA R	AGA R	<b>G1G</b> >	61G V	ဦ ဝ	<b>5</b> %	GA	<
1519	AAA	Noti RAATGAGGGGGCGC K	Ž	Noti Seco	8															•	:

SUBSTITUTE SHEET (RULE 26)

LAI AV3

FIGURE 8A

FIGURE 8C FIGURE 8A FIGURE 8B

ATGGATGCAATGAAGAGAGGCTCTGCTGTGTGCTG NarI U

CTGCTGTGTGGAGCAGTCTTCGTTTCGCCCAGCCAGGAAATCCATGCCCGATTCAGAAGAGGCGCCAGAACA Signal cleavage O **K** ⋖ I H ш O ⋖ O U 37

tcagatgctaaagcatatgatacagaggtacataatgtttgggccacacatgtgttgcctgtgtacccacagaccc × > დ ჯ **>** > 37 181 aacccacaagaagtagtattggtaaatgtgacagaaaattttaacatgtggaaaaatgacatggtagaacag W Σ Δ Z × 3 Σ Z Ĺ, Z W

>

ပ

4

x

~

atgcatgaggatataatcagtttatgggatcaaagcctaaagccatgtgaaaattaaccccactctgtgtt H . × > U ۵, × H <u> AGTTTAAAGTGCACTGATTTGGGGAATGCTACTAATACCAATAGTAGTAATACCAATAGTAGTAGCGGGGAA</u> Ы ပ S S S z Z ഗ ഗ

Z က Z = H 0 > > Ω ы Z Z H > G Ω L S **>** > ~ > × Ы U K × S 109 253 85 325 109 397 133 61

FIGURE 8B

**ATGATGATGGAGAAAGGAGAGATAAAAAACTGCTCTTTCAATATCAGCACAAGCATAAGAGGTAAGGTGCAG GAAGAAGAGGTAGTAATTAGATCTGCCAATTTCACAGACAATGCTAAAACCATAATAGTACAGCTGAACCAA** tctgtaga/attaattgtacaggtgctggacattgtaacattagtagggca/aatggaatgccactttaaaa **AAAGAATATGCATTTTTTTATAACTTGATATACCAATAGATAATGATACTACCAGCTATACGTTGACA GCCCCGGCTGGTTTTGCGATTCTAAATGTAATAATAAGACGTTCAATGGAACAGGACCATGTACAAA**TGTC agttgtaacacctcattacacaggcctgtccaaaggtatcctttgagccaattcccatacattat 0 G G æ H -1 G H H **H**. တ -⊢ G H Δ M a ഗ × æ z Z æ H ഗ H Δ S بعا Z S z H > H H Ĺ × > z ۵, U ဟ Z ပ مم بعا H ပ z 四 Z z ပ æ O 0 4 × 4 , 0 × ပ H ဟ H × G **= ~** W 4 G U بعا ပ × معا > W M W W 829 469 685 229 253 277 613 757 901 157 205 301 541 181

## FIGURE 8C

**CAGATAGCTAGCAAATTAAGAGAACAATTTGGAAATAATAAAACAATAATC**TTTAAGCAATCCTCAGGAGGG **GGGTCCGAGATCTTCAGACCTGGAGGAGGAGATATGAGGGACAATTGGAGAAGTGAATTATATAAATATAAA** Gacccagaaattgtaacgcacagttttaattgtggaggggaatttttctactgtaattcaacacactgtt **CTCCCATGCAGAATAAAACAATTTATAAACATGTGGCAGGAAGTAGGAAAAGCAATGTATGCCCCTCCCATC AGCGGACAAATTAGATGTTCATCAAATATTTACAGGCTGCTATTAACAA**GAGATGGTGGTAATAACAA Z 0 S M æ H 0 S G ပ Σ Ы Z 4 **a** S ы . [→ × æ æ **~ U** H Z \* æ > z Z × H (L) Ω × 山 ഗ 4 H æ G ပ 0 × Σ O z ပ Ы 3 H ပ Δ X ပ H H ۵, z ပ z H بعا ഗ 4 U z v 0 3 U (L) S ĹĿ H O œ O S ۵, × S H H × z ပ α, H ¥ Н æ > L. لعا ω S H 4 a ပ <u>سَ</u> ы × NotI G ^ ^ Δ. Δ 973 325 1045 349 1189 1333 1405 1117 373 397 421 445 469 1477 1261

18/42

	A	U	(D	F	Æ	<b>«</b>	~	:
	5 K	E H	ဖွဲ့ ဖ	X	ည္က 🏊	<b>6</b> 1	₹≍	Z z
	A AGA	GAAA	TA1	ું હું ∢	ACC	ATG	GT3	ACT T
	ATG	CAG O	TAT	AAAG K	ညည	Z	ig O	D F
FIGURE 9A FIGURE 9B FIGURE 9C	ATGGATGCAATGAAGAGA M D A M K R	<b>8</b>	GTC >	GCT.	CA C	AAT	<u>ن</u> م	MAT
CR	SAT O	ည္သမ	A F	CAT	Ž.	<b>3</b> ×	<b>Š</b> ×	T
FIG	¥ Y Y	<u>ප</u> ස	51C.	S S S	ည္တည္	ပ်ပည် 🗷	¥ 1	CT.
	~	CTGCTGTGTGCTGCTGTGTGGAGCAGTCTTCGTTTCGCCCAGCCAG	Nafi Catgcccgattcagaagaggcagagtagaaaagttgtgggtcacagtctattatggg H A R F R R G A R V E K L W V T V Y Y G	A Signal cleavage TGTGGAAAGAAGCAACCACTCTATTTTGTGCATCAGATGCTA	AGGTACATAATGTTTGGGCCACACATGCCTGTGTACCCACAGACCCCA E V H N V W A T H A C V P T D P	M H	ည်လ	N
		IIC F	116	9ava 1616 0	ဦ်ပ	J <sub>z</sub>	30	TG.
		GIC	Z ×	FITT	ည်	FTT	SATC D	SATG
FIGURE 9A		A GCA	<b>A</b>	gnal CTA:	CAT( H	CAT	ပ္ပ် ⊁	<b>8</b> ×
RE		<b>8</b> 50	GTA	ACT T	A P	SA E	rta.	ည္သံပ
ายเ		រក្ស	AGA M	<b>4</b> Ω H	ည် 🗸	A P	AGT	Z
正		CTG •	Nari 2060 P A	A C	मु द्र	STA.	F F	Ž1
		CTC P	<sup>ႜႍ</sup> ႘ၟ ၑ	ည် 🔻	57.7 >	Z Z	MIN H	CT
		CIG F	AGA M	S a	ZZ	E a	SATU	STT
		GTG V	AGA R	Ž×	CAT	ri R	SAG	ည်ပ
		រក្សា	TIC	1 3 3 3 3 3	STA V	STA	50	DE 1
		ည်	CGAT	6 7 8	S E	STA	Σ Σ	D C
		CIC	်ပ္ပိ 🗸	CCT P	Ž t	E ZA	30	T T
	Δν3	666CT	CATGC H A	Signal cleavage GTACCTGTGGAAAGAAGCAACCACCACTCTATTTTGTGCATCAGATGCTAAAGCATAT V P V W K E A T T T L F C A S D A K A Y	GATACAGAGGTACATAATGTTTGGGCCACACATGCCTGTGTACCCACAGACCCAACCCA D T E V H N V W A T H A C V P T D P N P	CAAGAAGTAGTATTGGAAAATGTAACAGAACATTTTAACATGTGGAAAAATAACATGGTA Q E V V L E N V T E H F N M W K N N M V	GANCAGATGCAGGAGATATAATCAGTTTATGGGATCAAAGCCTAAAGCCATGTGTAAAA E Q M Q E D'I I S L W D Q S L K P C V K	TTAACCCCACTCTGTGTTACTTTAAATTGCAAGGATGTGAATGCTACTAATACCACTAAT  L T P L C V T L N C K D V N A T N T T N
	JR-FL 1	19	79	139	199	259 87	319	379

439	GATA	300	30	NY.	8	រិតិ	AGA	GAC	40	AAS	ATA	4 4 4		) E	Ę	Ė	6	Ë		į	
4	D S E G T M E R G E I K N C S F N I T T	S	ω	(D	_	Σ.	E)	<b>~</b>	G	ы	<b>H</b>	<u> </u>	ŠZ	နှို့ ပ	S	ر سا به ۱	ξz	Į H	ۆ <b>⊢</b>		æ
499	AGCA	IM	3AG	VTG/	9	ည်	AGA	AAC	A.	TAT	بال	##J.	i i	į	<b>4</b>	<u>E</u>	(	Ę	8	Ö	
	S I R D E V O K E Y A L F Y K L D V P		<u>س</u>	~	-	<b>&gt;</b>	0	×	ш	<b>&gt;</b>	<b>A</b>	-	<u></u>	<b>Y</b>	₹ ×	ָרָבָּ בָּי	5 0	5 >	>		ď
S	ATAG	AT	AT.	(T.N.	TAC	Ş	GCT	'ATJ	Ċ	TTC	ATA	10	Į.	ر و	(	ر پ	Ę.	£ .	į	Š	
187	I D N N T S Y R L I S C D T S V I T Q	~	~ ~	~	_	f-i	S	<b>&gt;</b> -	æ		<b>—</b>	် လ	ပြု	ğa		် က	5 >	₹ H	} 	<u>ş</u> 0	2
	SCCIC	3TCC	3	IGA1	AT(	S	TTG	Ų.	Ş	ATT	ນ	ATA	CAT	TAT	TCT	. <u> </u>	٠ ر	ູ້ວ	ניט	<b>.</b>	F
207	A C P K I S F E P I P I H Y C A P A G F	<u> </u>	<b></b>			· (A)	<u>[a,</u>	W	۵,	<b>—</b>	۵.	н	Ξ	<b>&gt;</b>	ပ	<b>*</b>	ָ מַלְּיָלָ מָלְ	<b>*</b>	် ရ	<u>.</u>	4
619	GCGATTCTAAAGTGTAATGATAAGACGTTCAATGGAAAAGGACCATGTAAAAATGTCAGC	rrcı	3	GTG	Z I	VIG.	ATA	<b>A</b> G	S	TTC	AAT	<b>4</b> 55	***	<b>₹</b> 99	A C C	TGI	3	<b>*</b>	GIC	\(\frac{1}{2}\)	£ 1
~	<b>«</b>		<u>بر</u>	<b>U</b>	~	7	۵	×	H	(a,	Z	ပ	×	O	Δ.	ر ي	×	z	>	တ	)
ന	ACAG	CAC	<b>LAT</b> G	TAC	Ş	ATG.	₹	TT	ष्ट	S	GTA	GTA	TCA TCA	ACT	₹	CTG	CTG	CI	<b>₹</b>	8	()
247	TVQCTHGIRPVVSTQLLLNG	<b>5</b>	٠ ح			- 	ပ	<b>—</b>	œ	٥.	>	>	S	H	0	-1	-	-	Z	ၒ	
	AGTCTAGCAGAAGAGAGGTAGTAATTAGATCTGACAATTTCACGAACAATGCTAAAACC	S S	.AGA	75	705	9	TAG	3	TT	AGA	TCI	GAC	AAT	TTC	ACG	25	AAT	ပ္ပ	3	S	()
267	S	~ .3	M	Ш		f-3	>	>	<b>H</b>	Œ	ဟ	۵	Z	(a,	H	Z	Z	K	×	H	
2	ATAA	ragi	ATAGTACAGCTGAAAGAATCTGTAGAAATTAAT <b>TGTACAGGTGCTGGAC</b> ATTGTAAC	ည်	3	200	MAT	CTG	TA(	38	ATT	AAT	Ę	ថ្ន	53	Scr	g	3	TGI	<b>₹</b>	( )
287	<b>H</b>	<i>ح</i> ر	0	<b>H</b>	_		ы	ဟ	>	ш	н	Z	U	H	O	4	O	<b>E</b>	ပ	Z	
919	ATTA	TAG	AGTAGAGCAAAATGGAATGACACTTTAAAACAGATAGTTATAALATTAAGAGAACAA	A S	ATC	35	ATG	Ž	Ç	ITA	\$	CAG	ATA	GTT	ATA	Z.A	TTA	AGA	3	3	-
307	<b>H</b>	-	~	<b>×</b>		~ ~	z	Ω	H		×	0	H	>	<b>H</b>	×	H	æ	M	0	

## FIGURE 9C

FIGURE 10A

FIGURE 10A

FIGURE 10C FIGURE 10B

LAI AV3-CD4-

**ATGGATGCAATGAAGAGAGGCTCTGCTGTGTGCTG** ≥ ⋖ A ≥

CTGCTGTGTGGAGCAGTCTTCGTTTCGCCCAGCAAATCCATGCCCGATTCAGAAGAGGGGCCAGAACA NarI ⋖ I H ш O **(3)** م ဟ Þ Þ **∀** 0 U 37

Signal cleavage 🛦 109 37

× > G > > H > 3 H ×

TCAGATGCTAAAGCATATGATACAGAGGTACATAATGTTTGGGCCACACATGCCTGTGTACCCACAGACCCC U Z × M 181 61

**AACCCACAGAGTAGTATTGGTAAATGTGACAGAAATTTTAACATGTGGAAAAATGACATGGTAGAACA** X, Ω × X Z Z Ы Z > -1 M O 253 85

atgcatgaggatataatcagtttatgggatcaaagcctaaagccatgtbaaattaaccccactctgtgt H × > ပ ۵, × 1 S O Δ -1 S H 325 109

<u> AGTTTAAAGTGCACTGATTTGGGGAATGCTACTAATACCAATAGTAGTAATACCAATAGTAGTAGCGGGGAA</u> Ш G S S ဟ z H Z ഗ S Z H Z H K Z ၒ ... Ω H ပ × -1 ഗ 397 133

FIGURE 10B

### 23/42

atgatgatgagaaaagataaaaactgctctttcaatatggcacaagcataagggtaaggtgcag **aaagaatatgcattttttataaa**cttgatataataccaatagatatactaccagctatacgttgaca GCCCCGGCTGGTTTTGCGATTCTAAATGTAATAATAAGACGTTCAATGGAACAGGACCATGTACAAATGTC <u> agttgtaacacctcagtcattacacaggcctgtccaaaggtatcctttgagccaattcccatacattattgt</u> **agcacagtacaatgtacacatggaa**ttaggccagtagtatcaactgctgctgttgaatggcagtctagca GAAGAAGAGGTAGTAATTAGATCTGCCAATTTCACAGACAATGCTAAAACCATAATAGTACAGCTGAACCAA tctgtagaaattaat<del>tgtacaggtgctggacattgt</del>aacattagtaggcaaatggaatgccactttaaa **CAGATAGCTAGCAAATTAAGAGAACAATTTGGAAATAATAAAACAATAATCTTTAAGCAA**TCCTCAGGAGGG ပ × ပ ഗ ပ O 0 Z ᆸ G H H ı H <u>வ</u> H × H Ω ပ H 0 ഗ Z Z × Δ တ بعا H K S H > H z × × Ω > ۵. × z Z > م z H S H ပ Z z ပ م ပ H Ĺ 叫 G U ۵ K æ Z, z U H × H 0 ⋖ 0 J ပ ۲ × S U W **— -**¥ ы œ ~ 4 H G H ပ Ś × > 2 × 0 G M S Ľ W U Σ Σ 469 157 541 685 229 829 613 205 253 277 901 301 973 325 181 757

×

**~** 

W

S

æ

3

z

0

æ

E

۵

G

G

G

۵,

æ

Ĺ,

Н

(d)

S

G

445

### 24/42

GACCCAGAAATTGTAACGCACAGTTTTAATTGTGGAGGGGAATTTTTCTACTGTAATTCAACACACTGTTT aatagtacttggtttaatagtacttggagtactgaaggccaaataacacygaaggaagtacacaatcaca **CTCCCATGCAGAATAAAACAATTTATAAACATGGTGCAGGAAGTAGGAAAAGCAATGTATGCCCCTCCCATC AGCGGACAAATTAGATGTTCATCAAATA**TTACAGGGCTGCTATTAACAAGAGATGGTGGTAATAACAAC **GGGTCCGAGATCTTCAGA**CCTGGAGGAGGAGATATGAGGGACAATTGGAGNAGTGAATTATATAAATNTAAA Z Z z 4 ၒ တ U Σ ပ K W Ω × œ G H Z z ה ה ה > တ Ы W G 0 ပ > O ပ M ပ Σ H H z z ഗ H H (a, 3 Z S ທ H بعا X ഗ 0 S × H ပ Z, H æ > H œ ø ပ G 1045 349 1117 1189 397 1333 1261 421

0 > æ æ × × ۵, 4 > ပ H ۵, Ы × > > 1405 469

Noti

1447 6096006

SUBSTITUTE SHEET (RULE 26)

FIGURE 10C

FIGURE 11A

40	;	×
DAA		¥
TO TO TO TO A TO A TO A TO A TO A TO A		Z
	, .	⋖
` <b>4</b> 'U		Ω
4	•	2

ATGGATGCAATGAAGAGA M D A M K R	GGGCTCTGCTGTGTGTGTGGAGCAGTCTTCGTTTCGCCCAGCCAG	CATGCCCGATTCAGAAGAGGCGAGAGTAGAAAAGTTGTGGGGTCACAGTCTATTATGGG	GTACCTGTGTGGAAAGAAGCAACCACCACTCTATTTTGTGCATCAGATGCTAAAGCAATAT  V P V W K E A T T T L F C A S D A K A Y	GATACAGAGGTACATAATGTTTGGGCCACACATGCCTGTGTACCCACGAGCCCAACCCA D T E V H N V W A T H A C V P T D P N P	CAAGAAGTAGTATTGGAAAATGTAACAGAACATTTTAACATGTGGAAAAATAACATGGTA Q E V V L E N V T E H F N M W K N N M V	GAACAGATGCAGGAGATATAATCAGTTTATGGGATCAAAGCCTAAAGCCATGTGTAAAA E Q M Q E D I I S L W D Q S L K P C V K	TTAACCCCACTCTGTGTTACTTTAAATTGCAAGGATGTGAATGCTACTAATACCACTAAT
	GIGCIG V L	AGAAGA R R	AAÄGAA K E	CATAAT H N	TTGGAA L E	GAGGAT E D	TGTGTT
	TGCTGT	CGATTC R F	GTGTGG V W	GAGGTA	GTAGTA V V	ATGCAG M 0	CCACTC
JR-FL AV3-CD4 <sup>-</sup> 1	666CTC	CATGCC H A	GTACCT V P	GATACA D T	CAAGAA Q E	GAACAG	TTAACC
JR-FL 4 1 1	19	79	139	199	259	319	379

	)		3	ì
į				•
		L	1	
	)	ĺ		•
			)	١
		,	)	ì
	•		)	

439	439 GATAGCGAGCGAACGATGGAGAGGAGAAATAAAAAACTGCTCTTTCAATATCACCACA 147 D S E G T M E R G E I K N C S F N I T T	TTCAATATCACCACA F N I T T
499	99	TTGATGTAGTACCA
559	5.9 7.8	CAGTCATACACAG S V I T Q
619	19	SCCCGGCTGGTTTT A P A G F
679 227	79	GTAAAATGTCAGC C K N V S
739	39	TUCTGCTAAATGGC L L L N G
799	99	ACAATGCTAAAACC N N A K T
859 287	<b>S B</b>	CTGGACATTGTAAC
919	61 70	ARITAAGAGAACAA K L R E Q

## FIGURE 110

ž :	TTTGAGAATAAAACAATAGTCTTTAATCACTCCTCAGGGGGGGCCCCAGAAATTGTAATG F E N K T I V F N H S S G G D P E I V M	N	<b>3</b> ×	A F	ATA	offo >	777	¥ z	TCA H	CTC	CTC S	99	<b>8</b>	999	D Q	<b>A</b>	E B	<b>4</b>	1G1 V	\$	E X
S		CACAGILILATIGIGGAGGAGTITITCTACTGTAATICAACACAGACTGTITAATAGT H S F N C G G E F F Y C N S T Q L F N S		ည် ၁	နှိ ပြွ	e S S	<b>8</b> 9	111 F	rii F	XIX X	ig o	A Z	E S	Ž	E CAC	<b>3</b> 0-	DI J	TT F	\{\frac{1}{2}} \text{z}	TAC	H
9 9 <b>3</b>	Œ	ACTIGGAATAATACIGAAGGGICAAATAACACIGAAGGAAATACTAICACACICCCA T w n n n t e g s n n t e g n t i t l b	Z Z	Z Z	Ď.	ξω	9	ညီ လ	Z Z	Z Z	T T	IGA E	<b>8</b> €	<b>3</b>	ATA	-71 H	) I	એ દુ	ACT L	S I	<b>4</b>
GCAGA C R	<b>«</b>	tgcagaataaaaattataaacatg <del>gtg</del> caggaagtaggaaaagcaatgtatgcccct c r i k <u>o</u> i i n m <b>v</b> o e v g k n m y a p	<b>3</b> ×	30	LITA	ATA I	Z z	ATG	<b>§</b> >	S o	E E	AGT	g o	3 -	<b>₹</b> ×	2 4	X Z	TAT *	ည်	S T	H
H H	Æ	CCCATCAGAGACAAATTAGATGTTCATCAAATATTACAGGGCTGCTATTAACAAGAGAT PIRGQIRLT RSSNITG LILTRD	GAC G	30	I	AGA R	1GT C	ည်လ	ST. S	Z Z	H	TAC	S C	ijŢ.	ည္သ	TA 1	A i	Ž t	§ <del>%</del>	AGA	Ħ,
GI	K	GGTGGTATTAATGAGAATGGGACCGAGATCTTCAGACCTGGAGGAGGAGATATGAGGGAC	ATG	EAG	Z	999	ACC T	GAG	ATC	F	AG.	باري م	မ္မဌ	\(\frac{1}{2}\)	SAG S	<b>9</b> 5	SAT	A TO	3 <b>%</b> R	35	y a
1 ₹	ပ	AATTGGAGAAGTGAATTATAAATATAAAGTAGTAAAATTGAACCATTAGGAGTAGCA N W R S E L Y K Y K V V K I E P L G V A	AGT	Ž S	L	ATA Y	ZX X	ATA Y	X X	<b>A</b> G1	YG.	3	₹ ~	1110	GAACC E P	ပ္ပြဲ 🏊 🛓	ET.	9	SAG	TAG V	Y V V
S H	4	CCCACCAAGGAAGAGTGGTGCAAAGAGAAAAATGAGCGGCCGC	<b>%</b> <	A GA	<b>8</b> 8	NGAC R	ऽम् ४	STG	<b>₹</b> ○	AGA A	S a	₹×	ATG -	ğ	8	28	Q				

FIGURE 12A

LAI CD4

FIGURE 12C FIGURE 12B FIGURE 12A

**ATGGATGCAATGAAGAGAGGGCTCTGCTGT** ပ tPA signal sequence . O Σ Ω

Nari G **~** ⋖ Ŧ H ш O ⋖ O U

agaacagaaaaattgtgggtcacagtctattatggggtacctgtgtggaagbaagaaaccaccacttttt H K × > ۵, > ပ > H > 3 ... × Ы

Signal cleavage

tgtgcatcagatgctaaagcatatgatacagaggtacataatgtttgggccacacatgcctgtgtacccaca ပ **« =** H 4 3 > z = > M H Ω **>** 4 × ဟ 181

gaccccaacccacaggagtagtattggtaatgtgacagaaaattttaacatggaaaaatgacatggta E Z Z E E > Z > H > > (L) 0 ۵, Z ۵, Ω 253 85 GAACAGATGCATGAGGATATAATCAGTTTATGGGATCAAAGCCTAAAGCCATGTGTAAAATTAACCCCACTC 325

×

1

S

0

0

S

Ш

L

109

TGTGTTAGTTTAAAGTGCACTGATTTGGGGAATGCTACTAATACCAATAGT'AGTAATACCAATAGTAG'TAGC S 397 133

S

S

Z

Z

H

G

H

Δ

ပ

,,

S

SUBSTITUTE SHEET (RULE

109

### FIGURE 12B

GGGGAAATGATGATGGAGAAAGGAGAGATAAAAACTGCTCTTTCAATATCAGCACAAGCATAAGAGGTAAG **GTGCAGAAAGAATATGCATTTTTTTTATAAA**CTTGATATAATACCAATAGAT&ATGATACTACCAGCTATACG CTAGCAGAAGAAGAGGTAGTAATTAGATCTGCCAATTTCACAGACAATGCTAAAAACCATAATAGTACAĞCTG TTGACAAGTTGTAACACCTCAGTCATTACACAGGCCTGTCCAAAGGTATCCTTTGAGCCAATTCCCATACAT **TATIGIGCCCCGGCTGGTTTTGCGATTCTAAAATGTAATAATAAGACGTTCAATGGAACAGGACCAT**GTACA **aaccaatctgtagaaa**ttaattgtacaagacccaacaatacaagaaaaaagtatccgtatccagagggga **CCAGGGAGAGCATTTGTTACAATAGGAAAATAGGAAATATGAGACAAGCACATTGTAACATTAGTA**GAGCA ഗ 0 S G æ Z , G ۲ (L) ပ Ω 0 × **⊼** .. ഗ بعا 4 S H **a**. Z S Z > O Δ × × > H ۵, Œ H > Z م Z ഗ Z z (Le U ىم U z ပ Z æ z Z Ω 4 G **«** 0 × × H بم H G S × H æ æ H ω **>** H × K H G ပ [a, ပ × > S Ĺ. Z Ĺų ၒ 0 > (L) ~ ᆸ X M E M Ш ပ ဟ × S M S Ľ ပ 0 G > 205 829 973 325 469 157 541 181 613 685 229 757 253 277 901 301

29/42

# FIGURE 12C

**AAATGGAATGCCACTTTAAAACAGATAGCTAGCAAATTAAGAGAACAATTTGGAAATAATAAAACAAT**£ATC ¥ z Z ပ بعا 0 (L) æ \_\_ × ဟ K H 0 × . × 1045 349

TTTAAGCAATCCTCAGGAGGGACCCAGAAATTGTAACGCACAGTTTTAATTGTGGAGGGGAATTTTTTCTAC S I H > H Ы D, **A** G ပ S S 0

> 1117 373

ပ

ပ

z

Ĺ

tgtaattcaacacaactgtttaatagtacttggtttaatagtacttggagtäctgaagggtcaaataacact 1189 397

回 S 3 H S Z Ĺų 3 H S z Ĺ. ... 0 ဟ

GAAGGAAGTGACACACACTCCCATGCAGAATAAAACAATTAAAACATGGTGAGGAAGTAGGA:AA ပ Σ z 0 × ∝. ပ .-H H H Δ ဟ ပ M 1261 421

30/42

GCAATGTATGCCCCTCCCATCAGCGGACAAATTAGATGTTCATCAAATATTACAGGGCTGCTATTAACAAGA H .. ပ H H Z တ S ပ **~** H 0 ပ S ρ, 1333 445

GATGGTGGTAATAACAACAATGGGTCCGAGATCTTCAGACCTGGAGGAGGAGATATGAGGGACAATTGGAGA æ æ E Ω G ၒ ၒ Δ, œ Ĺ H Ы S G Z Z Z Z Ć G Ω 1405 469

4 > ပ H ۵, ഥ  $\vdash$ × > > × × H 1477 493

· NotI

**GTGGTGCAGAGAAAAATGAGOGGCCGC** 1549

Ы

æ

0

517

ATGGATGCAATGAAGAGA

FIGURE 13A

FIGURE 13A FIGURE 13B FIGURE 13C FIGURE 13D

GGGCTCTGCTGTGTGCTGCTGTGTGTGAGCAGTCTTCGTTTC&CCCAGCCAGGAAATC ш Σ O ⋖ ဟ Ω . .≅ ۵ ဟ Þ ⋖ 0 U r r Nari H Þ ບ ບ H G

CATGCCCGATTCAGAAGAGGCGCAGAGTAGAAAAGTTGTGGGTCACAGTCTATTATGGG G × H > **7** 山 > × 4 O × ட I 79

Signal cleavage

gtacctgtgtggaaagaagcaaccaccacttattttgtgcatcagatgctaaagcatat S H H Ы ×

139

GITACAGAGGTACATAATGTTTGGGCCACACATGCCTGTGTACCCACAGACCCCAACCCA > ပ Z L 199 67 CAAGAAGTAGTATTGGAAAATGTAACAGAACATTTTAACATGTGGAAAAATAACATGGTA × X z W Z 259 87

SUBSTITUTE SHEET (RULE 26)

19

JR-FL CD4"

### FIGURE 13E

* ¥	Z Z	Ž t	<del>ပ</del> ္ပဲ	950	
E O M O E D I I S L W D O S L K P C V K	ACCCCACTCTGTGTTAAATTGCAAGGATGTGAATGCTACTAATACCACTAAT T P L C V T L N C K D V N A T N T T N	AGCGAGGGAACGATGGAGAGGAGAATAAAAAACTGCTCTTTCAATATCACCACA S E G T M E R G E I K N C S F N I T T	ATAAGAGATGAGGAGAAAAATATGCTCTTTTTTATAAAC1'TGATGTAGTACCA I R D E V Q K E Y A L F Y K L D V V P	ATAGATAATAATAATACCAGCTATAGGTTGATAAGTTGTGACACCTCAGTCATTACACAG I D N N T S Y R L I S C D T S V I T Q	TGTCCAAAGATATCCTTTGAGCCAATTCCCATACATTATTGTGCCCCGGCTGGTTTT C P K I S F E P I P I H Y C A P A G F
រក្សា	<b>A</b> CC	ATC	GTA V	ATT	<b>6</b> €7
ర్ట్	AAT	AAT	GAT	GTC V	တ္ပ် ႕
<b>₹</b>	ACT	- H H H H H	Cit	TCA s	ည်
CTA	<b>.</b> 6€1	TCT	\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\	ACC T	<b>161</b> 0
ည် နှ	Z Z	ပ္ပို့ပ	TAT	<b>28</b> 0	TAT
ဦ ဝ	1GTG V	250	TTT	7GT	CAT
<b>A</b>	<b>1</b>	<b>₹</b> ×	CTI	<b>.</b> \$G1	IATA
YIQ X X	X ×	Į,	ည်	SATA I	ည္
H	ညီ ဂ	ξ E	X X	STTC L	Z-
S S	32	) (3)	AGA. E	<b>5</b> ₹2	ည္တ
Ž T	III	SAG R	ZX ×	X	IGA(
IAI	TAC	Ε Ε	S <sub>o</sub>	CAG S	E L
455 2	161 V	SAT M	661 <	TAC T	ATC. S
452 A	<u>ဂ</u> ဂ	AC T	1GA	A Z	GAT. I
50	ACT	ဗ္ဗဗ္ဗ	<b>8</b>	A Z	\$×
E	ပ္တ	<b>5</b>	28	A Z	700 P
5 a	<b>3</b> F	TAG	5	<b>3</b>	តិវិទ
ξ ω	TTN 1	GAT	S S	AI	GCC3 ▼
107	379 127	439	167	559 187	619

FIGURE 130

ACAGTACAATGTACACATGGAATTAGGCCAGTAGTATCAACTGCTGCTAAATGGC  T V Q C T H G I R P V V S T Q L L L N G  T V Q C T H G I R P V V S T Q L L L N G  AGTCTAGCAGAAGAAGAAGAATTAGATCTGACAATTTCACGAACAATGCTAAAACC  S L A E E V V I R S D N F T N N A K T  S L A E E V V I R S D N F T N N A K T  B S ATAATAGTACAGCTGAAAATCTGTAGAATTAATTGTACAAGACCCAACAACAATACA  I I V Q L K E S V E I N C T R F N N N T  B D AGAAAAAGAAAATAGAACCAGGGAGAAATTAATACTACAGGAAATAATAGGA  GATATAAGACAAAGCACATTGTAACATTAGTAGAACAAATGGAATTAAAACAG  GATATAAGACAAAGCACATTGTAACATTAGTAGAACAAATGGAAATGACACTTTAAAACAG  D I R Q A H C N I S R A K W N D T L K Q	227	3 ~	r i	7 7	\ <u>\</u>		Ž z	<u>ქ</u> ი	ž×	۲ ۲	TT.	<b>∑</b> 2	<u>წ</u> ი	<b>3×</b>	رقي د	ပ္ပြဲ မ	11G1	₹,	\\ \\ \\ \\ \\ \\ \\ \\ \\ \\ \\ \\ \\	STO \$	S .
24 96 88 40 42 4 97 97 97 97	8	ACA	GTA						: [TA]	. <b>V</b>	. ເວັ	֓֞֜֜֜֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓	ָרָלָי פַּבָּלָי	֝֞֜֜֜֜֜֜֜֜֝֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓	ָּ ק	ָּ ֭֭֭֭֓֞֞֞֝	ر رو	ن ا ا	ء آ <del>د</del> را		າ ປູ
or or or		g-s	>	O	ပ	<b>€</b> →	=	ပ	<b>H</b>	<b>«</b>	3 04	>	>	က	Ş <b>↔</b>	9	, ,	, .a		ŞZ	၌ ပ
r	66	AGT	Ct	Ş	3	3	CAC	25	513	LTZ	LY CY	MCI	SAC	:X	TIC	<b>A</b> CG	. <b>X</b>	**	ည်	₹	) <b>≯</b>
0 L 0 L 0 L	<u>[</u> 2	S	H	~	M	ы	(m)	>	>	H	æ	လ	Ω	z	<b>(4,</b>	H	<b>'Z</b> -	Z	~	×	H
- or or	<b>6</b>	ATA	ATA	GTA	ğ	CTC	3	3	5	CIA	3	דדא	TAT	TGI	NCA.	AGA		3	**	Ž	2
<b>0</b> L <b>0</b> L	22	<b>H</b>	<b>H</b>	>	ø	ı	×	M	S	>	M	H	Z	ပ	H	Œ	(L)	Z	Z	Z	H
L 0L		AGE	3	1987	IAT	3	CAT	000	Ş	9	SAG	Ş	LTI	TAT	.ACT	, AC	250	3	MT	ZI X	AGG,
91		æ	×	S	H	æ	H	G	Ω,	<b>U</b>	æ	4	(e,	<b>&gt;</b>	H	H	ප	ы	H	H	O
1		GAT	ATA	454	3	$\mathbf{g}$	3	rtgi	3	:AT	rAG1	FAGA	Ş	3	TG	:AA1	GAC	SAC:	ITT	₹	2
		Δ	H	æ	0	<	=	ပ	Z	H	S	æ	<	×	3	z	2	⊢	ı	×	0

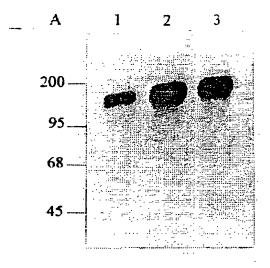
FIGURE 13D

1039	ATASI	TAT	3	LT S	ZX.	SAG	S	ATT	TGA	GAA	3	S	ATA	GIC	TIT	AAT	CAC	T.	AUL.
347	IVIKLREOFENKTIVFNHSS	<b>H</b>	<b></b>	<b>—</b>		~		<u>н</u>	M	Z	×	H	H	>	) )	z	I	Ś	က
1099	GGAGG	<b>49</b> 5	Ŝ	AGA:	3	TTG	[AA]	Ş	CAG.	TTT	CAA	FTGT	<b>4</b> 00	400	. 4	+++	<u> </u>	Į.	
367	G G D P E I V M H S F N C G G E F F Y C	Ω	Δ,	щ				<b>=</b>	GDPEIVMHSFNCGGEFFYC	Ĺ.	Z	ပ	ပ	<b>5</b>	<b>S</b> 🖼 .	<u> </u>	, L	<u>د</u> >-	
1159	AATTC	MC	Ş	<b>1</b> 20	5	35	TAG	TAC	TT	GAAT	Z.	[AA1	ACT	GAA		10°	AAT	YAA.	TO A
387	NSTOLFNSTERSTOTEGSNAT	H		-		~~ [:-	<b>U</b> )	<b>H</b>	3	Z	Z	z	H	ഥ	<b>ී</b>	S	z	z	H
1219	GAAGGAAATACTATCACACTCCCATGCAGAATAAAACAAATTATAAACATGGTGCAGGAA	A S	TAC	TAI	Ş	Š	ည်	ATG	CAG	AATA	3	Z Z	ATT	ATA	<b>A</b>	ATG	210	Ş	SGA
407	ம	2	<b>H</b>	-			. 3	ပ ၁	G N T I T L P C R I K Q I I N M V Q E	H	×	0	H	H	Z	Σ	<b>&gt;</b>	Ø	ы
1279	GTAGG	*	<b>₩</b>	X	CI)	त्र व	SS	TCC	CAT	CAG	35	3	ATT	AGA	TGT	Ţ	T C	3	LAT
427	V G K A M Y A P P I R G Q I R C S S N I	*	4	2,	~·		- 124	Δ,	<b>H</b>	<b>c</b> <	U	0	H	æ	U	ဟ	တ	Z	H
1339	ACAGG	GCT	S	ATT	ĭ. X	Ž	SAGA	TGG	<b>TGG</b>	<b>TAT</b>	Z	GAG	AAT	ည	SC	GAG	ATC	H	:AGA
447	TGLLTRDGGINENGTEIFR	 	<b>H</b>	<b>–</b>		<u></u>	~	<b>U</b> ,	U	<b>H</b>	Z	Ħ	Z	O	Ħ	ഥ	H	(m,	æ
1399	CCTGGAGGAGGAGATATGAGGGACAATTGGAGAAGTGAATTATATAAATATAAAGTAGTA	<b>308</b>	<b>8</b> 66	AGA	TAT	3	199	CA	TIG	SAGA	MG	3	TTA	TAT	X	TAT	3	GT	MGT
467	<b>9</b>	G	U		2,		~	Z	G G G D M R D N W R S E L Y K Y K V V	Œ	လ	M	-1	<b>&gt;</b>	×	<b>&gt;</b>	×	>	>
1459	AAAATTGAACCATTAGGAGTAGCACCCACCAAGGCAAAGAGAAGAGTGGTGCAAAGAGAA	TGA	ည	ATT	<b>3</b> 66	AG	NGC.	ACC	S S	<b>3</b> 86	Ş	<b>LAA</b> G	AGA	AGA	GTG	GTG	Ş	AG	NG &
. д	X	ы	Δ.	-1	G		<b>A</b>	٠.	EPLGVAPTKAKRVVOR	×	<b>«</b>	×	æ	æ	>	>	ø	æ	M
•		- 3	Noti	8															
1519	AAATGAGGGGGGG	3	3	3															
507	×																		

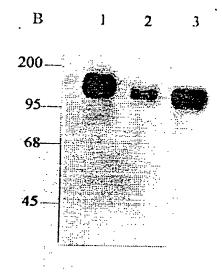
PCT/US94/03282

### 35/42

### FIGURE 14A



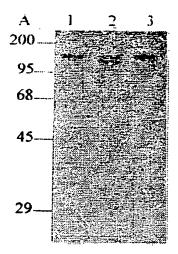
### FIGURE 14B



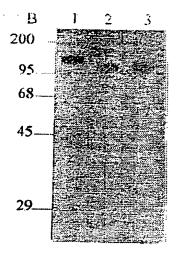
PCT/US94/03282

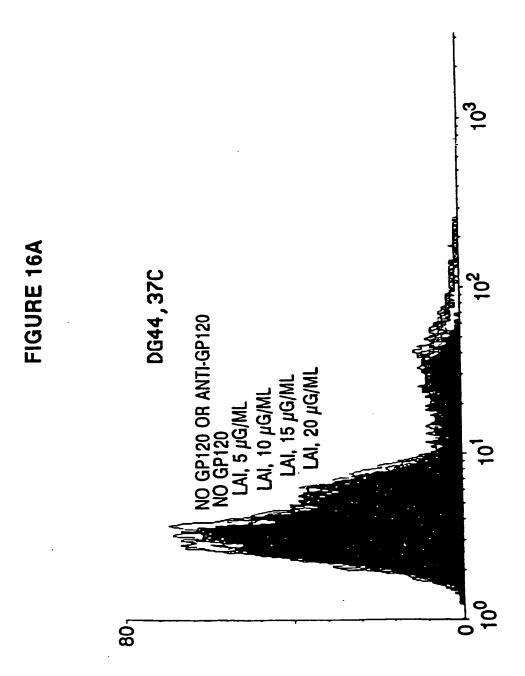
### 37/42

### FIGURE 15A

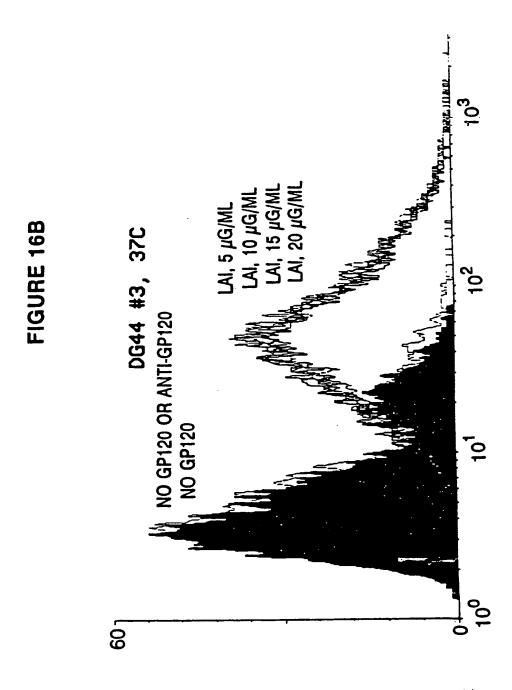


### FIGURE 15B

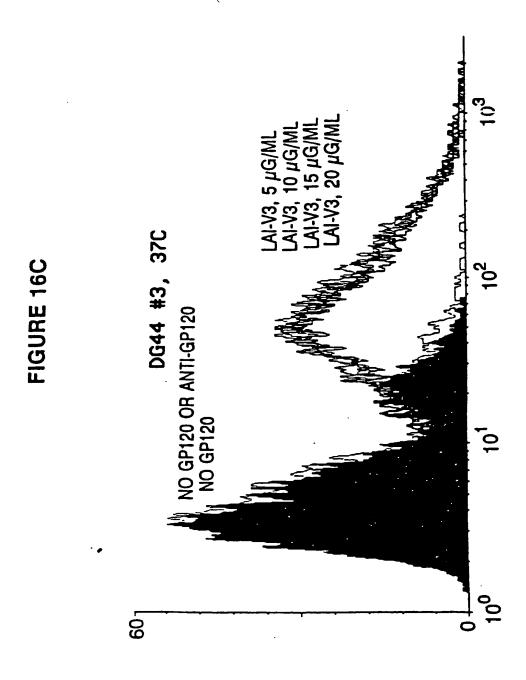




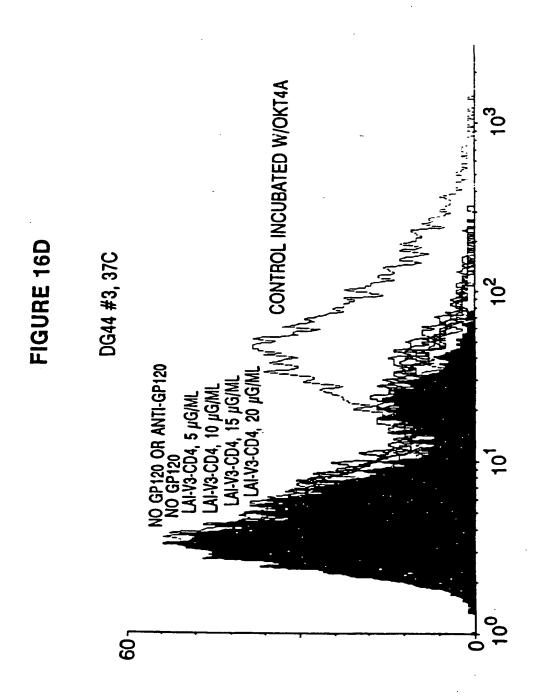
SUBSTITUTE SHEET (RULE 26)



SUBSTITUTE SHEET (RULE 26)



SUBSTITUTE SHEET (RULE 26)



SUBSTITUTE SHEET (RULE 26)

### INTERNATIONAL SEARCH REPORT

International application No. PCT/US94/03282

(A)	ACCIDICATION OF CIRIECT MARKET	
	ASSIFICATION OF SUBJECT MATTER :Please See Exira Sheet.	
	: 42A/88, 89; 536/27; 530/395	
	to International Patent Classification (IPC) or to both national classification and IPC	
	LDS SEARCHED	
wannam	documentation searched (classification system followed by classification symbols)	
U.S. :	424/88, 89; 536/27; 530/395	:
Document	tion searched other than minimum documentation to the extent that such documents are included	
	and the metadox	in the fields searched
Flectmaic	tota hase consulted during the investigation	
Liceronic	iata base consulted during the international search (name of data base and, where practicable	, search terms used)
APS, Di	alog, search terms: HIV-1, mutation, V3 loop, C4 region, envelope glycoprotein, v	accines musicinative
C. DOC	UMENTS CONSIDERED TO BE RELEVANT	
Category*	Citation of document with the	
Calegory	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No
Y	Science Volume 250	
'	Science, Volume 252, issued 17 May 1991, S. Wain-	6.7
1	HODSON, et al. "LAV Revisited: Origins of the Farly LIV 1	•
j	isolates from institut Pasteur", pages 961-965, see entire	
-	article.	
1		
Y	US, A, 5,030,449 (BERZOFSKY ET AL) 09 July 1991, cols.	
- 1	4-7.	1-27
- 1	• • • • • • • • • • • • • • • • • • • •	
Y	WO A 01/15510 (000000)	
_	WO, A, 91/15512 (GREGORY ET AL) 17 October 1991,	1-27
1	entire patent.	
. 1.		
Y	WO, A, 91/11461 (PASEK ET AL) 08 August 1991, see	1 77
1 •	entire patent.	1-27
1		
1	j	
ł		
i i		
- 1		
Further	documents are listed in the continuation of Box C. See patent family appear	
	categories of cited documents:  The later document published after the inter-	several filing date or priority
to be of	as defining the general state of the art which is not considered class and not in conflict with the application particular relevance principle or theory underlying the inven-	
earlier (	ocument published on or after the international filing data "X" document of particular extensions the	-4.1
docum-	considered povel or cannot be considered	to savelve so inventive step
	establish the publication date of another castion of other	•
4	document of particular relevance; the	
ED COLUMN		
	t muhlished seine to the international Climaton	EA .
docume		
- Prior	my date charmed	muy
- Prior	ny date charact of the same patent for	
e of the actua	ny date charact of the same patent for	
- Prior	of the international search  Date of mailing of the international search	
JUNE 1994	of the international search  Date of mailing of the international search  S. address of the ISA/IIS	
JUNE 1994 se and mailin	of the international search  Date of mailing of the international search  E address of the ISA/US	h report
JUNE 1994 te and mailing	of the international search  Date of mailing of the international search  E address of the ISA/US	h report
JUNE 1994 the and mailing manissioner of a PCT tahington, D.C.	of the international search  Date of mailing of the international search  E address of the ISA/US	=

### INTERNATIONAL SEARCH REPORT

International application No. PCT/US94/03282

A. CLASSIFICATION OF SUBJECT MATTER: IPC (5):	
A61K 39/12, 39/00; CO7K 17/00, 3/00, 13/00, 15/00; CO7H 15/12, 13/00	
;	
·	
· · · · · · · · · · · · · · · · · · ·	

### This Page is Inserted by IFW Indexing and Scanning Operations and is not part of the Official Record

### **BEST AVAILABLE IMAGES**

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

BLACK BORDERS

IMAGE CUT OFF AT TOP, BOTTOM OR SIDES

FADED TEXT OR DRAWING

BLURRED OR ILLEGIBLE TEXT OR DRAWING

SKEWED/SLANTED IMAGES

COLOR OR BLACK AND WHITE PHOTOGRAPHS

GRAY SCALE DOCUMENTS

LINES OR MARKS ON ORIGINAL DOCUMENT

REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY

### IMAGES ARE BEST AVAILABLE COPY.

☐ OTHER:

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.